

100

**H
H
H
-
O
H

O
H
O

O
H
O
H**

3

2

2

4
5
6
7
8

100

**H
H
H
-
O
H

O
H
O

O
H
O
H**

3

2

8

ALIGNMENTS

RESULT 1
 US-09-616-289-46
 ; Sequence 46, Application US/09616289
 ; Patent No. 6632923
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/09/616,289
 ; CURRENT FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 1638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1638)
 ; US-09-616-289-46

Query Match 100.0%; Score 1638; DB 4; Length 1638;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAGAACCAAGACAAAAGAACGGGGCTGCCAAATCCAAATCCAAAGACGCCCA	60
DB	1	ATGAAGAACCAAGACAAAAGAACGGGGCTGCCAAATCCAAATCCAAAGACGCCCA	60
QY	61	GGACACCGGAAGCAGACCCGAGGAGGCCAGAGCGGCCAGCCAGCGCGCTCTGCA	120
DB	61	GGACACCGGAAGCAGACCCGAGGAGGCCAGAGCGGCCAGCCAGCGCGCTCTGCA	120
QY	121	GTAGAAGCAGAGGTCCCGCAGCAGCGCGGTCTCTCGAAGCCGAGGGTCTCAAGCC	180
DB	121	GTAGAAGCAGAGGTCCCGCAGCAGCGCGGTCTCTCGAAGCCGAGGGTCTCAAGCC	180
QY	181	AGAACGGCTCAGTCTGGGGCCCTTCTGTGATGCTCTGAGGAGTGAAGCCCACTGGA	240
DB	181	AGAACGGCTCAGTCTGGGGCCCTTCTGTGATGCTCTGAGGAGTGAAGCCCACTGGA	240
QY	241	GACATCTGAGCACATACTGTGTGCAATAAACCAGGGGGGCCCGCGAGGATGGGCA	300
DB	241	GACATCTGAGCACATACTGTGTGCAATAAACCAGGGGGGCCCGCGAGGATGGGCA	300
QY	301	CAGGGTGAAGCGGCTGAACCCGAGATGACAGAAAGTCCCGGACCTATGTCAGGAAT	360
DB	301	CAGGGTGAAGCGGCTGAACCCGAGATGACAGAAAGTCCCGGACCTATGTCAGGAAT	360
QY	361	GGGAGGCTCAACCAACTCCAGTAGTCAATGGAGAGAGGAACCTCTCCAGGGGATCCA	420
DB	361	GGGAGGCTCAACCAACTCCAGTAGTCAATGGAGAGAGGAACCTCTCCAGGGGATCCA	420
QY	421	AACAGAGAGATCCGGCAGTGAAGGTTCGGAGACCGGACCATCGAAGGCCACAG	480

DB	421	AACACAGAGAGATCCGGCAGAGTGACGAGTCCGAGACCGAGACCATCGAAGCCACAG	480
QY	481	GAGAAGAAAAAGCCAAAGGTTTGGGTAAAGAGATCAGGTTGCTGATGACAGATTGAAT	540
DB	481	GAGAAGAAAAAGCCAAAGGTTTGGGTAAAGAGATCAGGTTGCTGATGACAGATTGAAT	540
QY	541	ACTCTGAGTACCCACAGAGGAGAGTGGTCTCTGTGCAAGAGATGCTGAACCTGCTG	600
DB	541	ACTCTGAGTACCCACAGAGGAGAGTGGTCTCTGTGCAAGAGATGCTGAACCTGCTG	600
QY	601	GAGGAGCACCGGATTCACAGAGACGATGAGCTCTACAGAAAAAGCAGAGCCAGCTG	660
DB	601	GAGGAGCACCGGATTCACAGAGACGATGAGCTCTACAGAAAAAGCAGAGCCAGCTG	660
QY	661	GTGCAAGAGAGAGACCACTGCGCGGTGACACAGCAAGGCCCTCTGCGCCCGCAGCAAG	720
DB	661	GTGCAAGAGAGAGACCACTGCGCGGTGACACAGCAAGGCCCTCTGCGCCCGCAGCAAG	720
QY	721	CTTGAGAGCCTATGCCGTGAGTGCAGCGGCACAAACCGCTCTCAAGAGAGAGTGTG	780
DB	721	CTTGAGAGCCTATGCCGTGAGTGCAGCGGCACAAACCGCTCTCAAGAGAGAGTGTG	780
QY	781	CAGCGGCCCGGAGGAGGAGGAGCGCAAGAGGTGACCTCGCACCTTCAGAGGTGACA	840
DB	781	CAGCGGCCCGGAGGAGGAGGAGCGCAAGAGGTGACCTCGCACCTTCAGAGGTGACA	840
QY	841	CTGAATCACAATTGAGTGCAGATGGAACAGACAATGAGCGCAACTTCAAGTTCGCGCAA	900
DB	841	CTGAATCACAATTGAGTGCAGATGGAACAGACAATGAGCGCAACTTCAAGTTCGCGCAA	900
QY	901	GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTCAGCAGATGATGAGTCCGCGAG	960
DB	901	GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTCAGCAGATGATGAGTCCGCGAG	960
QY	961	GAGCATATCGACAAAGTCTTCAAAACAAGGACCTCAACAGCAGCTGTGGTGTGCAAG	1020
DB	961	GAGCATATCGACAAAGTCTTCAAAACAAGGACCTCAACAGCAGCTGTGGTGTGCAAG	1020
QY	1021	CTCAGACAGCCCGAGAGATGCTTAAGGAGCGAGAGCGGCCAGCCAGCGGGAGAGGAT	1080
DB	1021	CTCAGACAGCCCGAGAGATGCTTAAGGAGCGAGAGCGGCCAGCCAGCGGGAGAGGAT	1080
QY	1081	TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTTGAGCTGTGATGAAGCAGCAAG	1140
DB	1081	TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTTGAGCTGTGATGAAGCAGCAAG	1140
QY	1141	ACCCACCTGAAAGCAAGCTTGCCTATACAGAGAAAGTTTGAAGGATTCAGAGACACA	1200
DB	1141	ACCCACCTGAAAGCAAGCTTGCCTATACAGAGAAAGTTTGAAGGATTCAGAGACACA	1200
QY	1201	CTTTCCAAAAGCAGCAGGATTTCAACACATTCAGCAGAGATGGAAGAGATGACTAAG	1260
DB	1201	CTTTCCAAAAGCAGCAGGATTTCAACACATTCAGCAGAGATGGAAGAGATGACTAAG	1260
QY	1261	AAGATCAAGAGCTGGAGAAAGAAACCAATGATGATGATGATGATGATGATGATGATG	1320
DB	1261	AAGATCAAGAGCTGGAGAAAGAAACCAATGATGATGATGATGATGATGATGATGATG	1320
QY	1321	AAGCGCTCTGATGATGCTGAGGAGAAACAGTCCGGGATTAAGAACTTGAAGGGCTG	1380
DB	1321	AAGCGCTCTGATGATGCTGAGGAGAAACAGTCCGGGATTAAGAACTTGAAGGGCTG	1380
QY	1381	CAGGTAAATAATCCAAACCGCTGGAGAAAGCTGTGCGGGCACTGACAGACGCGCAATGAC	1440
DB	1381	CAGGTAAATAATCCAAACCGCTGGAGAAAGCTGTGCGGGCACTGACAGACGCGCAATGAC	1440
QY	1441	CTGAACAGAGGGTACAGGACCTGAGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC	1500
DB	1441	CTGAACAGAGGGTACAGGACCTGAGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC	1500
QY	1501	CCTGAGAGAGGCCAGAGGGGCTTGGGGCTCAAGCACCAGCTCCCCCAGGGTCAAGAA	1560
DB	1501	CCTGAGAGAGGCCAGAGGGGCTTGGGGCTCAAGCACCAGCTCCCCCAGGGTCAAGAA	1560

Qy	1007	TGTTGGATGCCAAGCTCCAGAGGCCCGAGAGATGCTAAAGGAGGCAGAAAGCGGCACC	1066
Db	961	TGTTGGATGCCAAGCTCCAGAGGCCCGAGAGATGCTAAAGGAGGCAGAAAGCGGCACC	1020
Qy	1067	AGCGGAGAAAGGATTTTCTCTGTAAGAGGCAGTAGAGTCCACAGAGGATGTGTGAGCTGA	1126
Db	1021	AGCGGAGAAAGGATTTTCTCTGTAAGAGGCAGTAGAGTCCACAGAGGATGTGTGAGCTGA	1080
Qy	1127	TGAAGCAGCAGAGAGACCCACTGTAGAGCAACAGCTTGCCCTATACACAGAGAAATTTGAGG	1186
Db	1081	TGAAGCAGCAGAGAGACCCACTGTAGAGCAACAGCTTGCCCTATACACAGAGAAATTTGAGG	1140
Qy	1187	AGTTCCAGAAACACTTTTCCAAAAGCAGCAGGTATTTCCACATTTCAAGCAGGAGATGG	1246
Db	1141	AGTTCCAGAAACACTTTTCCAAAAGCAGCAGGTATTTCCACATTTCAAGCAGGAGATGG	1200
Qy	1247	AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAAACCACTATGACCGTCCCGGT	1306
Db	1201	AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAAACCACTATGACCGTCCCGGT	1260
Qy	1307	GGGAGAGCAGCAACAAAGGCCCTGCTTCAGATGGCTGAGGAGAAAACAGTCCGGGATAAAG	1366
Db	1261	GGGAGAGCAGCAACAAAGGCCCTGCTTCAGATGGCTGAGGAGAAAACAGTCCGGGATAAAG	1320
Qy	1367	AACGTGAGGGCTGCAGGTAAAAATCAAAGCTGGAGAGAGCTGTGCCGGCAGCTGAGAA	1426
Db	1321	AACGTGAGGGCTGCAGGTAAAAATCAAAGCTGGAGAGAGCTGTGCCGGCAGCTGAGAA	1380
Qy	1427	CAGAGCCCAATGACCTCAAACAAGAGGGTACAGGACCTGAGTGTGTGGCCACAGGGCTCCC	1486
Db	1381	CAGAGCCCAATGACCTGAAAACAAGAGGGTACAGGACCTGAGTGTGTGGCCACAGGGCTCCC	1440
Qy	1487	TCACGTGACAGTGGCCCTTGAGAGGAGGCCAGAGGGGCTTGGGGCTCAAGACCCAGCTCCC	1546
Db	1441	TCACGTGACAGTGGCCCTTGAGAGGAGGCCAGAGGGGCTTGGGGCTCAAGACCCAGCTCCC	1500
Qy	1547	CCAGGGTCACAGAAGCGCCCTTGCTTACCCAGGAGCAGCCGAGCACAGAAGCATCAGGCCAGA	1606
Db	1501	CCAGGGTCACAGAAGCGCCCTTGCTTACCCAGGAGCAGCCGAGCACAGAAGCATCAGGCCAGA	1560
Qy	1607	CTGGGCTCTAAGAGCCCACTTCGCCAGGGCC	1638
Db	1561	CTGGGCTCTAAGAGCCCACTTCGCCAGGGCC	1592

RESULT 4

US-09-616-289-17
Sequence 17, Application US/09616289
Patent No. 6632923
GENERAL INFORMATION:
APPLICANT: Lees, Robert S.
APPLICANT: Lees, Ann M.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 4697

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17

Query Match      87.9%; Score 1439; DB 4; Length 4697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 47 CAAAAAGCAGCCAGGACAAACCGGAAGCAGAGACCCGAGGGAGGCCAGAGCGGCCAGGCC 106
Db 1 CAAAAGCAGCCAGGACAAACCGGAAGCAGAGACCCGAGGGAGGCCAGAGCGGCCAGGCC 60

Qy 107 AGCGGGCTCTCAGGTAGAACGAGAGGTCCCGCAGCAGCCAGGCTTCCTCGAGAGCGG 166
Db 61 AGCGGGCTCTCAGGTAGAACGAGAGGTCCCGCAGCAGCCAGGCTTCCTCGGAAGCGG 120

Qy 167 AGGGTGCTCAAGCCAGAAACCGGCTCAGTCTCGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA 226
Db 121 AGGGGGCTCAAGCCAGAAACCGGCTCAGTCTCGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA 180

Qy 227 GCGGCCAACTGGAGACATATCTGAGCACAATCTGTGTGGAACAATAACAGGGGGGCCCGG 286
Db 181 GCGGCCAACTGGAGACATATCTGAGCACAATCTGTGTGGAACAATAACAGGGGGGCCCGG 240

Qy 287 GCAGAGATGGGACACAGGGTGAGCCGGCTCAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 346
Db 241 GCAGAGATGGGACACAGGGTGAGCCGGCTCAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 300

Qy 347 ATGTGGCAAGGAATGGGAGCCCTGAACCAACTCCAGTAGTCAATGGAGAGAAAGAACCTT 406
Db 301 ATGTGGCAAGGAATGGGAGCCCTGAACCAACTCCAGTAGTCAATGGAGAGAAAGAACCTT 360

Qy 407 CCAAGGGGATCAAAACACAGAGAGATCCGCAGAGTGACGAGTCCGAGAGCCGAGACC 466
Db 361 CCAAGGGGATCAAAACACAGAGAGATCCGCAGAGTGACGAGTCCGAGAGCCGAGACC 420

Qy 467 ATCGAAGGCCACAGGAGAGAAAGAAAGGTTTGGGTAAAGGATCACGTGTGCTGA 526
Db 421 ATCGAAGGCCACAGGAGAGAAAGAAAGGTTTGGGNAAGGATCACGTGTGCTGA 480

Qy 527 TGCAGACATTGATATCTTGAGTATCCACAGAGAGAAAGCTGGCTGTCTGTGACAGAAAT 586
Db 481 TGCAGACATTGATATCTTGAGTATCCACAGAGAGAAAGCTGGCTGTCTGTGCAAGAAAT 540

Qy 587 ATGCTGAATGCTGGAGGAGCACCGGAATTCAGAGAGCAGATGAAGCTCTCTACAGAAA 646
Db 541 ATGCTGAATGCTGGAGGAGCACCGGAATTCAGAGAGCAGATGAAGCTCTCTACAGAAA 600

Qy 647 AGCAGAGCCAGCTGGTGCAAGAGAGAGCAACCTGGCGGTGAGACACAGAAAGCCGTCC 706
Db 601 AGCAGAGCCAGCTGGTGCAAGAGAGAGCAACCTGGCGGTGAGACACAGAAAGCCGTCC 660

Qy 707 TGSCCCGACAGCTTGAGAGCCTATGCCGTGAGCTGACGGCGCAACCGCTCCCTCA 766
Db 661 TGSCCCGACAGCTTGAGAGCCTATGCCGTGAGCTGACGGCGCAACCGCTCCCTCA 720

Qy 767 AGGAAGAAGGTGTGACGGGGCCCGGAGAGGAGAGACGCAAGAGAGTGAACCTTCG 826
Db 721 AGGAAGAAGGTGTGACGGGGCCCGGAGAGGAGAGAGACGCAAGAGAGTGAACCTTCG 780

Qy 827 ACTTCAGGTGACATGTAATGACATTCAGCTGCAGATGGAAACAGCAATATGAGGCGCACT 886
Db 781 ACTTCCAGGTGACATGTAATGACATTCAGCTGCAGATGGAAACAGCAATATGAGGCGCACT 840

Qy 887 CCAAGCTGCCCAAGAGAGAACTAGAGCTGGCTGAGAGGCTCAAGAAAGCTGATTGAGCAGT 946
Db 841 CCAAGCTGCCCAAGAGAGAACTAGAGCTGGCTGAGAGGCTCAAGAAAGCTGATTGAGCAGT 900

Qy 947 ATGAGCTGCCGAGGAGCATATCGAAGAAAGTCTTCAAAACACAGAGACCTTCAACAGCAGC 1006

```

901	ATGAGCTGGCGGAGGAGCATATCGACAAAGCTTTCAAACAAGAGCACTTCAACACGACG	960
1007	TGCTGGATGCCAAAGCTCCACGACGCCCAGGAGATGCTAAAGGAGGAGAGACGGCACC	1066
961	TGCTGGATGCCAAAGCTCCACGACGCCCAGGAGATGCTAAAGGAGGAGAGACGGCACC	1020
1067	AGCGGAGAGAGGATTTTCTCTGAAGAGGACGATAGAGTCCACAGAGGATGTGTGACGTGA	1126
1021	AGCGGAGAGAGGATTTTCTCTGAAGAGGACGATAGAGTCCACAGAGGATGTGTGACGTGA	1080
1127	TGAAGCAGCAAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATACACAGAAAGTTTGAGG	1186
1081	TGAAGCAGCAAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATACACAGAAAGTTTGAGG	1140
1187	AGTTCCAGAACACACTTTTCCAAAGCAGCGAGTATTACCACTTCAAGCAGGAGATGG	1246
1141	AGTTCCAGAACACACTTTTCCAAAGCAGCGAGTATTACCACTTCAAGCAGGAGATGG	1200
1247	AAAAGATGACTAAGAGATCAAGAACTGGAGAAAGAAACCAACATGTACCGGTCCCGGT	1306
1201	AAAAGATGACTAAGAGATCAAGAACTGGAGAAAGAAACCAACATGTACCGGTCCCGGT	1260
1307	GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAAAAACAGTCCGGGATAAAG	1366
1261	GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAAAAACAGTCCGGGATAAAG	1320
1367	AACGTGAGGCGCTCGAGGTAAAAATCCACGGCTGGAGAAAGCTGTGCGGGCACTGCAGA	1426
1321	AACGTGAGGCGCTCGAGGTAAAAATCCACGGCTGGAGAAAGCTGTGCGGGCACTGCAGA	1380
1427	CAGAGCGCAATGACCTGAAACAAGAGGATCAGGACCTGAGTGTGCTGGCCAGGGGTCCC	1486
1381	CAGAGCGCAATGACCTGAAACAAGAGGATCAGGACCTGAGTGTGCTGGCCAGGGGTCCC	1440
1487	TCACCTGACAGTGGCCCTTGAGAGGAGCCAGAGGGGCTTGGGGTCAAGCAACCAGTCCC	1546
1441	TCACCTGACAGTGGCCCTTGAGAGGAGCCAGAGGGGCTTGGGGTCAAGCAACCAGTCCC	1500
1547	CCAGGGTCACAGAAAGCGCCTTGCTACCCAGGACCCGAGCACAGAAAGCATCAGGCCAGA	1606
1501	CCAGGGTCACAGAAAGCGCCTTGCTACCCAGGACCCGAGCACAGAAAGCATCAGGCCAGA	1560
1607	CTGGGGCTCAAGAGCCCACTCCGCCAGGGGCC	1638
1561	CTGGGGCTCAAGAGCCCACTCCGCCAGGGGCC	1592

RESULT 5

RESULT 3
PCT-US94-01101-1/c

PC1-US94-01101-1/8
: Sequence 1. Application PC/TUS9401101

; sequence I, Application:
: GENERAL INFORMATION:; GENERAL INFO:
; APPLICANT:

APPLICANT:
APPLICANT:
APPLICANT:

APPLICANT:
APPLICANT:
APPLICANT:

APPLICANT:
APPLICANT:
APPLICANT:

APPLICANT:
APPLICANT:
APPLICANT:

APPLICANT:
APPLICANT:
APPLICANT:

APPLICANT: HT
TITLE OF INVENTION: HT

;	TITLE OF INVENTION:	HI
:	TITLE OF INVENTION:	FA

; TITLE OF INVENTION: FACTOR: INTERLEUKIN-14
 : NUMBER OF SEQUENCES: 8

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

CORRESPONDENCE: W.

ADDRESSEE: WOLF, GRE
STREET, 600 ATLANTIC

STREET: 600
CITY: BOSTON

CITY: BOSTON
STATE: MASSACHUSETTSSTATE: MASSACHUSETTS
COUNTRY: USA; COUNTRY: USA
; EID: 00010

ZIP: 02210

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible

;	APPLICATION NUMBER:	PCT/US94/01101	
;	FILING DATE:	FILED HEREWITH	
;	PRIOR APPLICATION DATA:		
;	APPLICATION NUMBER:	US 08/005,156	
;	FILING DATE:	15-JAN-1993	
;	ATTORNEY/AGENT INFORMATION:		
;	NAME:	GATES, EDWARD R.	
;	REGISTRATION NUMBER:	31,616	
;	REFERENCE/DOCKET NUMBER:	B0819/7000WO	
;	TELECOMMUNICATION INFORMATION:		
;	TELEPHONE:	617/720-3500	
;	TELEFAX:	617/720-2441	
;	INFORMATION FOR SEQ ID NO: 1:		
;	SEQUENCE CHARACTERISTICS:		
;	LENGTH:	1854 base pairs	
;	TYPE:	nucleic acid	
;	STRANDEDNESS:	single	
;	TOPOLOGY:	linear	
;	MOLECULE TYPE:	cdna to mRNA	
;	HYPOTHEICAL:	NO	
;	ANTI-SENSE:	NO	
;	ORIGINAL SOURCE:		
;	ORGANISM:	HOMO SAPIENS	
;	CELL TYPE:	LYMPHOMA	
;	PCT-US94-01101-1		
	Query Match	47.3%; Score 774; DB 5; Length 1854;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 774; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	749	GGCACAAACCGCTCCCTCCAAAGGAAGAGGTGTGCAGCGGCCCGGGAGGAGGAGGAAGC	808
Db	1672	GGCACAAACCGCTCCCTCAAGGAGAGAGGTGTGCAGCGGCCCGGGAGGAGGAGGAAGC	1613
Qy	809	GCAGAGAGGTGACCTCGCACTTCAGGTGACACTGAAATGACATTGAGCTGCAGATGGAC	868
Db	1612	GCAGAGAGGTGACCTCGCACTTCAGGTGACACTGAAATGACATTGAGCTGCAGATGGAC	1553
Qy	869	AGCACAAATGAGCGCAACTCCAAGCTCGGCAAGAGAAACATGAGCTGGCTGAGAGGCTCA	928
Db	1552	AGCACAAATGAGCGCAACTCCAAGCTCGGCAAGAGAAACATGAGCTGGCTGAGAGGCTCA	1493
Qy	929	AGAAGCTGATTGACGAGTATGAGCTCGGAGAGAGCATATCGACAAAGTCTTCAAACACA	988
Db	1492	AGAAGCTGATTGACGAGTATGAGCTCGGAGAGAGCATATCGACAAAGTCTTCAAACACA	1433
Qy	989	AGGACCTACACAGCAGCTGGTGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGG	1048
Db	1432	AGGACCTACACAGCAGCTGGTGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGG	1373
Qy	1049	AGGCAGAAGAGCGGCACCAAGCGGGAAGGATTTTCTCTCTGAAAGGAGGAGTAGTCCC	1108
Db	1372	AGGCAGAAGAGCGGCACCAAGCGGGAAGGATTTTCTCTCTGAAAGGAGGAGTAGTCCC	1313
Qy	1109	AGAGGATGTGTGAGCTGTGATGAGCAGCACAAGACCACCTGAGACCAAGCTTGCCTAT	1168
Db	1312	AGAGGATGTGTGAGCTGTGATGAGCAGCACAAGACCACCTGAGACCAAGCTTGCCTAT	1253
Qy	1169	ACACAGAGAAGTTTGAGGAGTTCCAGAACACACTTTCCAAAAGCAGCAGGATATTCAACA	1228
Db	1252	ACACAGAGAAGTTTGAGGAGTTCCAGAACACACTTTCCAAAAGCAGCAGGATATTCAACA	1193
Qy	1229	CATTCAAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGCTGGAAGAAAGAACCA	1288
Db	1192	CATTCAAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGCTGGAAGAAAGAACCA	1133
Qy	1289	CCATGTACCGGTCCCGTGGGAGAGCAGCAACAAGGCCCTGCTTGGATGGCTGAGGAGA	1348
Db	1132	CCATGTACCGGTCCCGTGGGAGAGCAGCAACAAGGCCCTGCTTGGATGGCTGAGGAGA	1073
Qy	1349	AAACAGTCCGGGATAAAGAACTGAGGGGCTCCAGGTAAAAATCCAAACGGCTGGAGAGC	1408
Db	1072	AAACAGTCCGGGATAAAGAACTGAGGGGCTCCAGGTAAAAATCCAAACGGCTGGAGAGC	1013

QY 1409 TGTGCGGGCACTGAGACAGAGCGCAATGACCTGAAACAAGAGGTACAGACCTGAGTG 1468
DB 1012 TGTGCGGGCACTGAGACAGAGCGCAATGACCTGAAACAAGAGGTACAGACCTGAGTG 953
QY 1469 CTGGTGGCCAGAGCTCCCTCATCTGACAGTGGCCCTGAGAGGAGCCAGAGGGGC 1522
DB 952 CTGGTGGCCAGAGCTCCCTCATCTGACAGTGGCCCTGAGAGGAGCCAGAGGGGC 899

RESULT 6

US-09-616-289-51
; Sequence 51, Application US/09516289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-51

Query Match 20.5%; Score 336; DB 4; Length 22255;
Best Local Similarity 100.0%; Pred. No. 7.8e-148;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGAGCCG 231
DB 2999 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGAGCCG 3059
QY 232 CAACTGGAAGACATCTGAGCACATCTGTGGCAATAACAGGGGGGGCCCCGGCGAG 291
DB 3059 CAACTGGAAGACATCTGAGCACATCTGTGGCAATAACAGGGGGGGCCCCGGCGAG 3118
QY 292 GATGGGCGACAGGTGAGCCGGCTGAACCCGAGATGAGAGAGTCCCGGACCTATGTG 351
DB 3119 GATGGGCGACAGGTGAGCCGGCTGAACCCGAGATGAGAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCCTCAACCACTCCAGTAGTCAATGGAGAGAGCAACCTCCCAAG 411
DB 3179 GCAAGGAATGGGAGCCTCAACCACTCCAGTAGTCAATGGAGAGAGCAACCTCCCAAG 3238
QY 412 GGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGGAGCCGAGACCATCGA 471
DB 3239 GGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGGAGCCGAGACCATCGA 3298
QY 472 AGCCACAG 507
DB 3299 AGCCACAG 3334

RESULT 7

US-08-979-608A-14
; Sequence 14, Application US/08979608A

; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-Nov-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...1731
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-979-608A-14
Query Match 3.6%; Score 59; DB 4; Length 4722;
Best Local Similarity 100.0%; Pred. No. 5.6e-18;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1396 CGGTGAGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGG 1454
DB 1456 CGGTGAGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGG 1514
RESULT 8
US-09-517-849-14
; Sequence 14, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 309:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-309

Query Match
Best Local Similarity 1.8%; Score 30; DB 2; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1211 GCAGGAGGTATTCCACCATTCAAGCAGG 1240
Db 30 GCAGGAGGTATTCCACCATTCAAGCAGG 1

RESULT 12
US-08-859-998-310
Sequence 310, Application US/08859998
Patent No. 5994076
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Jekhade, George
APPLICANT: Bibilashvilli, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,998
FILING DATE: 21-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-310

Query Match
Best Local Similarity 1.8%; Score 30; DB 2; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 860 AGATGGAACAGCACATGAGCGCAACTCCA 889
Db 1 AGATGGAACAGCACATGAGCGCAACTCCA 30

RESULT 13
US-09-225-928-309/c
Sequence 309, Application US/09225928
Patent No. 6352829
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Jekhade, George
APPLICANT: Bibilashvilli, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,928
FILING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,998
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 309:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
SEQUENCE DESCRIPTION: SEQ ID NO: 309:
US-09-225-928-309

Query Match
Best Local Similarity 1.8%; Score 30; DB 4; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1211 GCAGGAGGTATTCCACCATTCAAGCAGG 1240
Db 30 GCAGGAGGTATTCCACCATTCAAGCAGG 1

```

STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,201B
APPLYING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/859,998
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 309:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-09-225-201B-309
Query Match 1.8%; Score 30; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 30; Conservative 0; Mismatches 0; Indels 0
QY 1211 GCAGCGAGGTATTCCACCATTCAGACGAGG 1240
Db 30 GCAGCGAGGTATTCCACCATTCAGACGAGG 1
RESULT 16
US-09-225-201B-310
; Sequence 310, Application US/09225201B
; Patent No. 6489455
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
Jokhadze, George
Bibilashvili, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,201B
FILING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/859,998
FILING DATE: 21-MAY-1997

```


ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
SEQUENCE DESCRIPTION: SEQ ID NO: 310:
US-09-225-201B-310
Query Match 1.8%; Score 30; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 860 AGATGGAACAGACAAATGAGCGCAACTCCA 889
Db 1 AGATGGAACAGACAAATGAGCGCAACTCCA 30
RESULT 17
US-08-979-608A-42
Sequence 42, Application US/08979608A
Patent No. 6355451
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,608A
FILING DATE: 26-Nov-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,547
FILING DATE: 03-JUN-1997
APPLICATION NUMBER: US 60/031,930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...21
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-08-979-608A-42
Query Match 1.3%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGTCCTCTGAGGAGCTG 225
Db 1 CGTGATGTCCTCTGAGGAGCTG 21
RESULT 18
US-09-517-849-42
Sequence 42, Application US/09517849
Patent No. 6605588
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...21
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-09-517-849-42
Query Match 1.3%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGTCCTCTGAGGAGCTG 225
Db 1 CGTGATGTCCTCTGAGGAGCTG 21

RESULT 19
US-09-616-289-42
; Sequence 42, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-42

Query Match 1.3%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGTCCTCAGGAGCTG 225
DB 1 CGTGATGTCCTCAGGAGCTG 21

RESULT 20
US-09-220-132-64/c
; Sequence 64, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shvjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-64

Query Match 1.3%; Score 21; DB 4; Length 851;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGACGACG 147
DB 83 GCAGAGGTCCTCCGACGACG 63

RESULT 21
US-09-621-976-3527/c
; Sequence 3527, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PE2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3527
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 195..389
; NAME/KEY: misc.feature
; LOCATION: 457..497
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-3527

Query Match 1.2%; Score 20; DB 4; Length 614;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1235 AGCAGGAGATGGAAGATG 1254
DB 40 AGCAGGAGATGGAAGATG 21

RESULT 22
US-08-968-751-5
; Sequence 5, Application US/08968751
; Patent No. 5948643
; GENERAL INFORMATION:
; APPLICANT: Rubinfeld, Bonnie
; APPLICANT: Polakis, Paul G.
; APPLICANT: Ligenfelter, Carol
; APPLICANT: Vuong Terilyn T.
; TITLE OF INVENTION: MODULATORS OF BRCA1 ACTIVITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond
; STATE: CA
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/968,751
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX1024 GG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 262-8710
; TELEFAX: (510) 222-9758
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1191 base pairs

;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 34..1191
US-08-968-751-5

Query Match 1.2%; Score 20; DB 2; Length 1191;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCAGAGGAGAGCTGCG 569
|||||
Db 451 ACCCAGAGGAGAGCTGCG 470

RESULT 23

US-08-817-997A-1/c
; Sequence 1, Application US/08817997A
; Patent No. 5827719
; GENERAL INFORMATION:
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Sakari
; APPLICANT: Kofod, Lene V.
; TITLE OF INVENTION: An Enzyme With Lipolytic
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58277190 No. 5827719disk Of No. 5827719th America, Inc.
; STREET: 405 Lexington Avenue - 64th Fl.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,997A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 7435

ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4316.204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9652
; TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 994 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 43...729
; OTHER INFORMATION:

US-08-817-997A-1

Query Match 1.2%; Score 19; DB 1; Length 994;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1307 GCGAGAGCAGCAAGGC 1325
|||||
Db 95 GCGAGAGCAGCAAGGC 77

RESULT 24

US-09-328-352-2648
; Sequence 2648, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2648
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2648

Query Match 1.2%; Score 19; DB 4; Length 1209;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 ACATTCAAGCAGGAGATGG 1246
|||||
Db 1127 ACATTCAAGCAGGAGATGG 1145

RESULT 25

US-09-621-976-2162
; Sequence 2162, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2162
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..223
US-09-621-976-2162

Query Match 1.1%; Score 18; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1271 AGCTGGAGAAAGAACCA 1288
|||||
Db 291 AGCTGGAGAAAGAACCA 308

RESULT 26

US-09-621-976-9955/c
; Sequence 9955, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm
; SEQ ID NO 9955
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-9955

Query Match 1.1%; Score 18; DB 4; Length 479;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GAGAGGAGGCCAGAGGGG 1521
|||
Db 182 GAGAGGAGGCCAGAGGGG 165

RESULT 27
US-09-109-266-9/c
; Sequence 9, Application US/09109266
; Patent No. 5970111
; GENERAL INFORMATION:
; APPLICANT: Weng, Zude
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Klein, Theodore M.
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Cell Proliferation-Associated Proteins
; FILE REFERENCE: BS-1275-P1
; CURRENT FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 9
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Zea mays
US-09-109-266-9

Query Match 1.1%; Score 18; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 GGAGGAGGAGGAGGAGG 809
|||
Db 50 GGAGGAGGAGGAGGAGG 33

RESULT 28
US-09-252-991A-4286
; Sequence 4286, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4286
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4286

Query Match 1.1%; Score 18; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GCCCGCGCGAGGATGGG 298
|||
Db 122 GCCCGCGCGAGGATGGG 139

RESULT 29
US-08-998-416-413/c
; Sequence 413, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgene
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 413:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 753 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1307RP
US-08-998-416-413

Query Match 1.1%; Score 18; DB 3; Length 753;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1173 AGAAGGTTTGAGGAGTT 1190
|||
Db 222 AGAAGGTTTGAGGAGTT 205

RESULT 30
US-08-592-126-97
; Sequence 97, Application US/08592126

; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Septin-2.seq
US-08-592-126-97
Query Match 1.1%; Score 18; DB 1; Length 1534;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 797 AGGAGGAGGAGCGCAAGG 814
Db 1335 AGGAGGAGGAGCGCAAGG 1352
RESULT 31
US-09-168-595-97
Sequence 97, Application US/09168595
Patent No. 655666
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/168,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,126
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Septin-2.seq
US-09-168-595-97
Query Match 1.1%; Score 18; DB 4; Length 1534;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 797 AGGAGGAGGAGCGCAAGG 814
Db 1335 AGGAGGAGGAGCGCAAGG 1352
RESULT 32
US-09-277-565-16/c
Sequence 16, Application US/09277565
Patent No. 6207384
GENERAL INFORMATION:
APPLICANT: Metalanos, John J.
APPLICANT: Akerley, Brian J.
APPLICANT: Rubin, Eric J.
TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL
GENES BY IN VITRO TRANSDUCTION MUTAGENESIS
FILE REFERENCE: 00742/052002
CURRENT APPLICATION NUMBER: US/09/277,565
CURRENT FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: 60/079,770
EARLIER FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 1599
TYPE: DNA
ORGANISM: Haemophilus influenzae
US-09-277-565-16
Query Match 1.1%; Score 18; DB 3; Length 1599;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1221 ATTCAACACATTCACGCA 1238
Db 1117 ATTCAACACATTCACGCA 1100
RESULT 33
US-09-620-312D-948/c
Sequence 948, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
POLYPEPTIDES
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 948
LENGTH: 2391
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (668)...(1930)
US-09-620-312D-948

Query Match 1.1%; Score 18; DB 4; Length 2391;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GAGAGAGCCGAGGGG 1521
|||
Db 304 GAGAGAGCCGAGGGG 287

RESULT 34
US-08-895-601-2/c
Sequence 2, Application US/08895601
Patent No. 6060262
GENERAL INFORMATION:
APPLICANT: Beer-Romero, Peggy
APPLICANT: Strack, Peter J.
APPLICANT: Glass, Susan J.
APPLICANT: Rolfe, Mark
TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,
METHODS AND REAGENTS RELATED THERETO
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,601
FILING DATE: 16-JUL-1997

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-096.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2790 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2782
US-08-895-601-2

Query Match 1.1%; Score 18; DB 3; Length 2790;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 GGAGGAGGAGGAGGCG 809
|||
Db 22 GGAGGAGGAGGAGGCG 5

RESULT 35
US-09-702-953B-1/c
Sequence 1, Application US/09702953B
Patent No. 6673897
GENERAL INFORMATION:
APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
TITLE OF INVENTION: NOVEL INHIBITORS OF NF-kappaB ACTIVATION
FILE REFERENCE: 2676-4554US
CURRENT APPLICATION NUMBER: US/09/702,953B
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: PCT/BE99/000055
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: 98201472.2
PRIOR FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2812
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (117)..(2060)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (81)..(116)
OTHER INFORMATION:
US-09-702-953B-1

Query Match 1.1%; Score 18; DB 4; Length 2812;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 CTCAGTCTGGGGCCCTTC 205
|||
Db 2692 CTCAGTCTGGGGCCCTTC 2675

RESULT 36
US-09-620-312D-673/c
Sequence 673, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom

```

; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radojko T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 673_
; LENGTH: 2900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)...(2762)
; US-09-620-312D-673

Query Match 1.1%; Score 18; DB 4; Length 2900;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1364 AAGAACTGGAGGGCTGCG 1361
Db 1846 AAGAACTGGAGGGCTGCG 1829

RESULT 37
US-09-091-952A-6
; Sequence 6, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Deters-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1...8065
; OTHER INFORMATION: Clone 22
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116...1036
; OTHER INFORMATION: Clone 22 coding region
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 452...505
; OTHER INFORMATION: alternatively spliced portion
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5595...5685
; OTHER INFORMATION: amplified region for genotyping
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-091-952A-6

Query Match 1.1%; Score 18; DB 4; Length 8065;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1271 AGCTGGAGAAAGAAACCA 1288
Db 1057 AGCTGGAGAAAGAAACCA 1074

RESULT 38
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6378225
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(111282)

```

; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 1.1%; Score 18; DB 4; Length 111282;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1457 AGGACCTGAGTGTGCTG 1474
|||||
Db 41515 AGGACCTGAGTGTGCTG 41532

RESULT 39

US-09-214-808-1
; Sequence 1, Application US/09214808A
; Patent No. 6475793
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Andre
; APPLICANT: Freiberg, Christoph
; APPLICANT: Perret, Xavier Philippe
; APPLICANT: Broughton, William John
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; Patent No. 6475793
; TITLE OF INVENTION: Plasmid
; FILE REFERENCE: CARP0068
; CURRENT APPLICATION NUMBER: US/09/214,808A
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/IB97/00950
; PRIOR FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 536165
; TYPE: DNA
; ORGANISM: Rhizobium
US-09-214-808-1

Query Match 1.1%; Score 18; DB 4; Length 536165;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 865 GAACAGCAACATGAGCGC 882
|||||
Db 119889: GAACAGCAACATGAGCGC 119906

RESULT 40

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 1.1%; Score 18; DB 4; Length 1830121;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1221 ATTCACCATTCACGA 1238
|||||
Db 940375 ATTCACCATTCACGA 940358

RESULT 41

US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid


```

; APPLICATION NUMBER: GB 9603069.7
; FILING DATE: 14-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-452
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-133-321-9

Query Match      1.0%; Score 17; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels

QY 1308 GGAGAGCAGCAACAGG 1324
    |||
Db   61 GGAGAGCAGCAACAGG 77

RESULT 44
US-09-133-321-10/c
; Sequence 10, Application US/09133321
; Patent No. 6255558
; GENERAL INFORMATION:
; APPLICANT: Haseloff, Phillip J.
; APPLICANT: Hodge Sarah
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye, P.C.
; STREET: 1100 No. 6255558th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: VA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/133,321
; FILING DATE: 12-AUG-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: GB 9603069.7
; FILING DATE: 14-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-452
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-133-321-10

Query Match      1.0%; Score 17; DB 3; Length 83;

```

Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0;

QY 1308 GGAGAGCAGCAACAGG 1324
|||||
Db 83 GGAGAGCAGCAACAGG 67

RESULT 45
US-08-791-495-1
; Sequence 1, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-3441
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-791-495-1

Query Match 1.0%; Score 17; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ACCGAGGGAGCCGAG 94
|||||
Db 63 ACCGAGGGAGCCGAG 79

RESULT 46
US-09-328-111-809/c
; Sequence 809, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 809
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(388)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-809

Query Match 1.0%; Score 17; DB 3; Length 388;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 CTCGAAGGAGAGGTTGT 779
|||||
Db 94 CTCGAAGGAGAGGTTGT 78

RESULT 47
US-09-702-705-1636
; Sequence 1636, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1636
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1636

Query Match 1.0%; Score 17; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 592 GAACGTCTGGAGGAGCA 608
|||||
Db 192 GAACGTCTGGAGGAGCA 208

RESULT 48
US-09-736-457-1636
; Sequence 1636, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1636
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1636
Query Match 1.0%; Score 17; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 592 GAACTGCTGGAGGACCA 608
DB 192 GAACTGCTGGAGGACCA 208
|||||
RESULT 49
US-09-614-124B-1636
; Sequence 1636, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1636
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-1636
Query Match 1.0%; Score 17; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 50
US-09-671-325-1636
; Sequence 1636, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1636
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1636
Query Match 1.0%; Score 17; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 592 GAACTGCTGGAGGACCA 608
DB 192 GAACTGCTGGAGGACCA 208
|||||
RESULT 51
US-09-621-976-152/c
; Sequence 152, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 152
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 71...496
; NAME/KEY: sig_peptide
; LOCATION: 71...121
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.59999990463257
; OTHER INFORMATION: seq AKLASWTSMALA/AS
US-09-621-976-152
Query Match 1.0%; Score 17; DB 4; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 715 AGCAAGCTTGAGGCCT 731
DB 97 AGCAAGCTTGAGGCCT 81
|||||

```
RESULT 52
US-09-621-976-16898
; Sequence 16898, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16898
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16898

Query Match 1.0%; Score 17; DB 4; Length 540;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1081 TTCTCTCTGAAGAGGC 1097
DB 495 TTCTCTCTGAAGAGGC 511

RESULT 53
US-08-469-427A-10/c
; Sequence 10, Application US/08469427A
; Patent No. 5607918
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Ulf
; APPLICANT: Olofsson, Birgitta
; APPLICANT: Alitalo, Kari
; APPLICANT: Pajusola, Katri
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
; TITLE OF INVENTION: DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,427A
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,651
; FILING DATE: 01-MAR-1995
; NAME: Evans, Joseph D
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 41979CP2
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: human fibrosarcoma
US-08-469-427A-10/c

Query Match 1.0%; Score 17; DB 2; Length 570;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1514 CAGAGGGGCTGGGGCT 1530
DB 428 CAGAGGGGCTGGGGCT 412

RESULT 54
US-08-609-443B-10/c
; Sequence 10, Application US/08609443B
; Patent No. 5840693
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf
; APPLICANT: OLOFSSON, Birgitta
; APPLICANT: ALITALO, Kari
; APPLICANT: PAJUSOLA, Katri
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
; TITLE OF INVENTION: DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan, P.L.L.C.
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,443B
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,651
; FILING DATE: 01-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/469,427
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/569,063
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/41979CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: human fibrosarcoma
US-08-609-443B-10/c

Query Match 1.0%; Score 17; DB 2; Length 570;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1514 CAGAGGGGCTGGGCT 1530
DB 428 CAGAGGGGCTGGGCT 412

RESULT 55
US-08-569-063C-10/c
; Sequence 10, Application US/08569063C
; Patent No. 5928939
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf
; APPLICANT: OLOFSSON, Birgitta
; APPLICANT: ALITALO, Kari
; APPLICANT: PAJUSOLA, Katri
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
; TITLE OF INVENTION: DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan, P.L.L.C.
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,063C
; FILING DATE: 06-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/469,427
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,651
; FILING DATE: 01-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/41979CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: human fibrosarcoma
; US-08-569-063C-10

Query Match 1.0%; Score 17; DB 2; Length 570;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1514 CAGAGGGGCTGGGCT 1530
DB 428 CAGAGGGGCTGGGCT 412

RESULT 56
US-08-851-896-10/c
; Sequence 10, Application US/08851896
; Patent No. 6331301
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf

QY 1514 CAGAGGGGCTGGGCT 1530
DB 428 CAGAGGGGCTGGGCT 412

RESULT 57
US-08-609-443B-14/c
; Sequence 14, Application US/08609443B
; Patent No. 5840893
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf
; APPLICANT: OLOFSSON, Birgitta
; APPLICANT: ALITALO, Kari
; APPLICANT: PAJUSOLA, Katri
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
; TITLE OF INVENTION: DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 57

QY 1514 CAGAGGGGCTGGGCT 1530
DB 428 CAGAGGGGCTGGGCT 412

Query Match 1.0%; Score 17; DB 4; Length 570;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1514 CAGAGGGGCTGGGCT 1530
DB 428 CAGAGGGGCTGGGCT 412

RESULT 58
US-08-851-896-10/c
; Sequence 10, Application US/08851896
; Patent No. 6331301
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,063C
FILING DATE: 06-DEC-1995
FOR APPLICATION DATA:

APPLICATION NUMBER: US 08/469,427
FILING DATE: 06-JUN-1995
IOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,651
FILING DATE: 01-MAR-1995
TORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D
REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/41979CP3
LECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 base pairs
TYPE: nucleic acid

STRANDNESS: single
TOPOLOGY: linear
LECULE TYPE: cDNA
IGINAL SOURCE:
TISSUE TYPE: human
69-063C-14
Match 1.0%

Local Similarity 100.0%;
 es 17; Conservative 0;

1514 CAGAGGGGCCTGGGCT 153
 |||||
 526 CAGAGGGGCCTGGGCT 510

59
51-896-14/c
nce 14, Application US/08851896
t No. 6331301
RAL INFORMATION:
PLICANT: ERIKSSON, Ulf
PLICANT: OLOFSSON, Birgitta
PLICANT: ALTALO, Kari

APPLICANT: PAJUSOLA, Katri
 TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
 TITLE OF INVENTION: DNA CODING THEREFOR
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Evenson, Mckewen, P.L.L.C.
 STREET: 1200 G Street, N.W., Suite 700
 CITY: Washington

STATE: DC COUNTRY: USA
Zip: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0. Version #1.25

RECENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,896
FILING DATE: 06-MAY-1997
IOR APPLICATION DATA:
APPLICATION NUMBER: US/08/609,443B
FILING DATE: 01-MAR-1996
IOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,651

FILING DATE: 01-MAR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/469,427
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/569,063
FILING DATE: 06-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/41979CP4
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
TISSUE TYPE: human
US-08-851-896-14

Query Match 1.0%; Score 17; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1514 CAGAGGGGCTGGGCT 1530
DB 526 CAGAGGGGCTGGGCT 510

RESULT 60
US-09-133-321-1
Sequence 1, Application US/09133321
Patent No. 625558
GENERAL INFORMATION:
APPLICANT: Haseloff, Phillip J.
APPLICANT: Hodge, Sarah
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye, P.C.
STREET: 1100 No. 625558th Glebe Road, 8th Floor
CITY: Arlington
STATE: VA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/133,321
FILING DATE: 12-AUG-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9603069.7
FILING DATE: 14-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C.
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 604-452
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: N/A
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 701 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 17..694
US-09-133-321-1

Query Match 1.0%; Score 17; DB 3; Length 701;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1308 GGAGAGCAGCAACAGG 1324
DB 418 GGAGAGCAGCAACAGG 434

RESULT 61
US-09-530-685A-30
Sequence 30, Application US/09530685A
Patent No. 6541619
GENERAL INFORMATION:
APPLICANT: DAEWOONG PHARMACEUTICAL CO., LTD. et al.
TITLE OF INVENTION: A FUSION PROTEIN OF HGF AND HUMAN ANGIOGENIN, AND PROCESS FOR
TITLE OF INVENTION: PREPARING THE SAME
FILE REFERENCES: P0012/DW/US
CURRENT APPLICATION NUMBER: US/09/530,685A
CURRENT FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: KR 1997-57603
PRIOR FILING DATE: 1997-11-01
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Kopatentin 1.71
SEQ ID NO 30
LENGTH: 717
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fusion gene of angiogenin and hEGF
US-09-530-685A-30

Query Match 1.0%; Score 17; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 787 GCCCGGGAGGAGGAGGA 803
DB 539 GCCCGGGAGGAGGAGGA 555

RESULT 62
US-08-896-164-40
Sequence 40, Application US/08896164
Patent No. 6218521
GENERAL INFORMATION:
APPLICANT: OBATA, Yuichi
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,164
FILING DATE: July 17, 1997

```

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 753 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-40

```

```

; Query Match 1.0%; Score 17; DB 3; Length 753;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 23 ACGGGGCTGCCAACAA 39
Db 313 ACGGGGCTGCCAACAA 329

```

```

RESULT 63
US-09-489-039A-5078
; Sequence 5078, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5078
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-5078

```

```

; Query Match 1.0%; Score 17; DB 4; Length 783;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1119 TGAGCTGATGAGCAGC 1135
Db 153 TGAGCTGATGAGCAGC 169

```

```

RESULT 64
US-09-530-685A-29
; Sequence 29, Application US/09530685A
; Patent No. 6541619
; GENERAL INFORMATION:
; APPLICANT: DAEMONG PHARMACEUTICAL CO., LTD. et al.
; TITLE OF INVENTION: A FUSION PROTEIN OF HGF AND HUMAN ANGIOGENIN, AND PROCESS FOR
; FILE REFERENCE: P0012/DW/US
; CURRENT APPLICATION NUMBER: US/09/530,685A
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: KR 1997-57603
; PRIOR FILING DATE: 1997-11-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 29
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: fusion gene of angiogenin and hEGF
; US-09-530-685A-29

```

```

; Query Match 1.0%; Score 17; DB 4; Length 927;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 787 GCCCGGAGGAGGAGGA 803
Db 365 GCCCGGAGGAGGAGGA 381

```

```

RESULT 65
US-08-353-400-24/c
; Sequence 24, Application US/08353400
; Patent No. 5665357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353.400
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324819.3
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411089.7
; FILING DATE: 03-JUN-1994
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 940 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-353-400-24

```

```

; Query Match 1.0%; Score 17; DB 1; Length 940;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 791 GGGAGGAGGAGGAGAG 807
Db 873 GGGAGGAGGAGGAGAG 857

```

```

RESULT 66
US-08-303-569B-4/c
; Sequence 4, Application US/08303569B
; Patent No. 5859205
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Emage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,569B
FILING DATE: 07-SEP-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 18..722
FEATURE: mat_peptide
LOCATION: 84..722
US-08-303-569B-4

Query Match 1.0%; Score 17; DB 2; Length 943;
Best Local Similarity 100.0%; Pred.No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 GGGAGGAGGAGGAGAG 807
DB 865 GGGAGGAGGAGGAGAG 849

RESULT 67
US-08-116-247-4/c
Sequence 4, Application US/08116247
Patent No. 5929212
GENERAL INFORMATION:
APPLICANT: Jolliffe, Linda K.
APPLICANT: Zivin, Robert A.
APPLICANT: Adair, John R.
APPLICANT: Athwal, Diljeet S.
TITLE OF INVENTION: CD3 Specific Recombinant Antibody
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5929212ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/116,247
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/743,377
FILING DATE: 10-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Paintin, Francis A.
REGISTRATION NUMBER: 19,386
REFERENCE/DOCKET NUMBER: CARP-0011
TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 18..722
US-08-116-247-4

Query Match 1.0%; Score 17; DB 2; Length 943;
Best Local Similarity 100.0%; Pred.No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 GGGAGGAGGAGGAGAG 807
DB 865 GGGAGGAGGAGGAGAG 849

RESULT 68
US-09-795-515-4/c
Sequence 4, Application US/09795515
Patent No. 6632927
GENERAL INFORMATION:
APPLICANT: Adair, John R.
APPLICANT: Athwal, Diljeet S.
APPLICANT: Emtage, John S.
TITLE OF INVENTION: Humanised Antibodies
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6632927ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/795,515
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,658
FILING DATE: 01-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0057
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 18..722
FEATURE: mat_peptide
LOCATION: 84..722
US-09-795-515-4

Query Match 1.0%; Score 17; DB 4; Length 943;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 GGGAGGAGGAGGAGAAG 807
 |||||
 Db 865 GGGAGGAGGAGGAGAAG 849

RESULT 69

US-08-791-495-4
 ; Sequence 4, Application US/08791495
 ; Patent No. 5811519

GENERAL INFORMATION:

APPLICANT: Leth, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Falleur, Thierry
 TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: L0461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1002 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 65..697
 US-08-791-495-4

Query Match 1.0%; Score 17; DB 1; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ACCGAGGAGGCCAGG 94
 |||||
 Db 625 ACCCGAGGAGGCCAGG 641

RESULT 70

US-09-252-991A-1988/c
 ; Sequence 1988, Application US/09252991A
 ; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 1988
 LENGTH: 1044
 TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1988

Query Match 1.0%; Score 17; DB 4; Length 1044;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 525 GATGCAGACATTGAATA 541
 |||||

Db 116 GATGCAGACATTGAATA 100

RESULT 71

US-08-933-750C-90
 ; Sequence 90, Application US/08933750C
 ; Patent No. 5932442

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Bandman, Olga
 APPLICANT: Shah, Purvi
 APPLICANT: Au-Young, Janice
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Corley, Neil C.
 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 NUMBER OF SEQUENCES: 98
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/933,750C
 FILING DATE: September 23, 1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0356 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 90:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1078 base pairs
 TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SININOT01
; CLONE: 2184712
US-08-933-750C-90

Query Match 1.0%; Score 17; DB 2; Length 1078;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 591 TGAAGTCTGGAGGAGC 607
DB 589 TGAAGTCTGGAGGAGC 605

RESULT 72
US-09-234-613-90
; Sequence 90, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SININOT01
; CLONE: 2184712
US-09-234-613-90

Query Match 1.0%; Score 17; DB 3; Length 1078;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 591 TGAAGTCTGGAGGAGC 607
DB 589 TGAAGTCTGGAGGAGC 605

RESULT 73
US-09-482-273-42/c
; Sequence 42, Application US/09482273
; Patent No. 6534831
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (582)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (592)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-482-273-42
```

```
Query Match 1.0%; Score 17; DB 4; Length 1164;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 780 GCAGCGGCGCGGAGG 796
DB 376 GCAGCGGCGCGGAGG 360
```

RESULT 74

```
US-08-948-705-7
; Sequence 7, Application US/08948705A
; Patent No. 6043084
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ASSOCIATED WITH COLON CANCER AND METHODS FOR DIAGNOSING AND
; FILE REFERENCE: LUD-5506-JEL/NDH
; CURRENT APPLICATION NUMBER: US/08/948,705A
; CURRENT FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-08-948-705-7
```

```
Query Match 1.0%; Score 17; DB 3; Length 1306;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
```

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 CGTGAGCTGCAGCGGCA 752
 Db 94 CGTGAGCTGCAGCGGCA 110

RESULT 75

US-09-510-543-7
 ; Sequence 7, Application US/09510543
 ; Patent No. 6517837
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Matthew J.
 ; APPLICANT: Chen, Yao-Tseng
 ; APPLICANT: Stockert, Elisabeth
 ; APPLICANT: Old, Lloyd J.
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH COLON CANCER AND
 ; FILE REFERENCE: LUD-5506.1-JEL/NDH
 ; CURRENT APPLICATION NUMBER: US/09/510,543
 ; CURRENT FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 09/102,322
 ; PRIOR FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 22
 ; SEQ ID NO 7
 ; LENGTH: 1306
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-510-543-7

Query Match 1.0%; Score 17; DB 4; Length 1306;

Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 CGTGAGCTGCAGCGGCA 752
 Db 94 CGTGAGCTGCAGCGGCA 110

RESULT 76

US-08-484-993B-17
 ; Sequence 17, Application US/08484993B
 ; Patent No. 5837497
 ; GENERAL INFORMATION:
 ; APPLICANT: Harris Ph.D., Jeffrey D.
 ; APPLICANT: Hsu, Kuang T.
 ; APPLICANT: Podolski, Joseph S.
 ; TITLE OF INVENTION: Materials and Methods for Immunocontraception
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/484,993B
 ; FILING DATE: 09-NOV-1993
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/012,990
 ; FILING DATE: 29-JAN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/973,341
 ; FILING DATE: 09-NOV-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clough, David W.

REGISTRATION NUMBER: 36,107
 REFERENCE/DOCKET NUMBER: 31745
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6653
 TELEFAX: 312/474-0448
 TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1319 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Felis domesticus
 DEVELOPMENTAL STAGE: Juvenile
 HAPLOTYPE: Diploidy
 TISSUE TYPE: Ovary
 CELL TYPE: Oocyte
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 26..1297
 US-08-484-993B-17

Query Match 1.0%; Score 17; DB 2; Length 1319;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGAGCTGGCT 570
 Db 516 CAGAGGAGAGCTGGCT 532

RESULT 77

US-08-484-158B-17
 ; Sequence 17, Application US/08484158B
 ; Patent No. 5976545
 ; GENERAL INFORMATION:
 ; APPLICANT: Harris Ph.D., Jeffrey D.
 ; APPLICANT: Hsu, Kuang T.
 ; APPLICANT: Podolski, Joseph S.
 ; TITLE OF INVENTION: Pharmaceutical Compositions for
 ; TITLE OF INVENTION: Immunocontraception
 ; NUMBER OF SEQUENCES: 61
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/484,158B
 ; FILING DATE: 07-JUNE-95
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/149,223
 ; FILING DATE: 09-NOV-93
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/012,990
 ; FILING DATE: 29-JAN-93
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/973,341
 ; FILING DATE: 09-NOV-92
 ; ATTORNEY/AGENT INFORMATION:

NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1319 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Felis domesticus

DEVELOPMENTAL STAGE: Juvenile

HAPLOTYPE: Diploidy

TISSUE TYPE: Ovary

CELL TYPE: Oocyte

FEATURE:

NAME/KEY: CDS

LOCATION: 26..1297

US-08-484-158B-17

Query Match

Best Local Similarity 1.0%; Score 17; DB 2; Length 1319;

Pred. No. 2.7e+02;

Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGAGCTGGCT 570

|||||

DB 516 CAGAGGAGAGCTGGCT 532

RESULT 78

US-08-484-596A-17

; Sequence 17, Application US/08484596A

; Patent No. 5981228

; GENERAL INFORMATION:

; APPLICANT: Harris Ph.D., Jeffrey D.

; APPLICANT: Hsu, Kuang T.

; APPLICANT: Podolski, Joseph S.

; TITLE OF INVENTION: Materials and Methods for Immunocontraception

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,596A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/149,223

; FILING DATE: 11-NOV-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/973,341

; FILING DATE: 09-NOV-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Clough, David W.

; REGISTRATION NUMBER: 36,107

; REFERENCE/DOCKET NUMBER: 31745

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1319 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Felis domesticus

DEVELOPMENTAL STAGE: Juvenile

HAPLOTYPE: Diploidy

TISSUE TYPE: Ovary

CELL TYPE: Oocyte

FEATURE:

NAME/KEY: CDS

LOCATION: 26..1297

US-08-484-596A-17

Query Match

Best Local Similarity 1.0%; Score 17; DB 2; Length 1319;

Pred. No. 2.7e+02;

Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGAGCTGGCT 570

|||||

DB 516 CAGAGGAGAGCTGGCT 532

RESULT 79

US-08-480-150A-17

; Sequence 17, Application US/08480150A

; Patent No. 5989550

; GENERAL INFORMATION:

; APPLICANT: Harris Ph.D., Jeffrey D.

; APPLICANT: Hsu, Kuang T.

; APPLICANT: Podolski, Joseph S.

; TITLE OF INVENTION: Materials and Methods for Immunocontraception

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/480,150A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/149,223

; FILING DATE: 09-NOV-1993

; APPLICATION NUMBER: 08/012,990

; FILING DATE: 29-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/973,341

; FILING DATE: 09-NOV-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Clough, David W.

; REGISTRATION NUMBER: 36,107

; REFERENCE/DOCKET NUMBER: 31745

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6653

; TELEFAX: 312/474-0448

```
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-480-150A-17

Query Match 1.0%; Score 17; DB 2; Length 1319;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGGAAGCTGGCT 570
DB 516 CAGAGGAGGAAGCTGGCT 532

RESULT 80
US-08-458-731-17
Sequence 17, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
```

```
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-458-731-17

Query Match 1.0%; Score 17; DB 3; Length 1319;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGGAAGCTGGCT 570
DB 516 CAGAGGAGGAAGCTGGCT 532

RESULT 81
US-08-149-223A-17
Sequence 17, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
```

HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploid
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-149-223A-17

Query Match 1.0%; Score 17; DB 3; Length 1319;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CAGAGGAGAGCTGGCT 570
Db 516 CAGAGGAGAGCTGGCT 532

RESULT 82
US-08-896-164-84
Sequence 84, Application US/08896164
Patent No. 6218521
GENERAL INFORMATION:
APPLICANT: OBATA, Yuichi
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
WITH GASTRIC CANCER AND METHODS FOR
TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,164
FILING DATE: July 17, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6218521man D. Hanson
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 1347 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-896-164-84

Query Match 1.0%; Score 17; DB 3; Length 1347;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 ACGGGGCTGCCAACA 39
Db 907 ACGGGGCTGCCAACA 923

RESULT 83

US-09-482-273-56/c
Sequence 56, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 56
LENGTH: 1559
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1445)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1551)
OTHER INFORMATION: n equals a,t,g, or c
US-09-482-273-56

Query Match 1.0%; Score 17; DB 4; Length 1559;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1019 AGTCCAGCAGGCCAG 1035
Db 392 AGTCCAGCAGGCCAG 376

RESULT 84
US-07-937-609-22
Sequence 22, Application US/07937609
Patent No. 5319073
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992

PRIOR APPLICATION DATA: US 07/928,033
APPLICATION NUMBER: 29,768
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1621 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: guinea pig gallbladder and pancreas CCKA
FEATURE:
NAME/KEY: CDS
LOCATION: 205..1497
US-07-937-609-22

Query Match 1.0%; Score 17; DB 1; Length 1621;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 790 CGGGAGGAGGAGGAGAA 806
Db 1538 CGGGAGGAGGAGGAGAA 1554

RESULT 85
US-08-029-170-22
Sequence 22, Application US/08029170
Patent No. 6163173
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1621 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: guinea pig gallbladder and pancreas CCKA
FEATURE:
NAME/KEY: CDS
LOCATION: 205..1497
US-08-029-170-22

Query Match 1.0%; Score 17; DB 3; Length 1621;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 790 CGGGAGGAGGAGGAGAA 806
Db 1538 CGGGAGGAGGAGGAGAA 1554

RESULT 86
US-08-683-743-3
Sequence 3, Application US/08683743 12/1/98
Patent No. 5843697
GENERAL INFORMATION:
APPLICANT: Pestka, Sidney
APPLICANT: Kotenko, Serguei
TITLE OF INVENTION: CYTOKINE RECEPTOR SIGNAL TRANSDUCTION
TITLE OF INVENTION: CHAIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,743
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1875 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-683-743-3

Query Match 1.0%; Score 17; DB 2; Length 1875;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 788 CCGGGAGGAGGAGGAG 804
Db 1532 CCGGGAGGAGGAGGAG 1548

RESULT 87

US-08-733-622C-1

; Sequence 1, Application US/08733622C

; Patent No. 6528634

; GENERAL INFORMATION:

; APPLICANT: Georgopoulos, Katia

; APPLICANT: Morgan, Bruce A.

; TITLE OF INVENTION: AIOLOS GENE

; FILE REFERENCE: 10287-030001

; CURRENT APPLICATION NUMBER: US/08/733,622C

; CURRENT FILING DATE: 1996-10-17

; PRIOR APPLICATION NUMBER: US 60/017,646

; PRIOR FILING DATE: 1998-05-14

; PRIOR APPLICATION NUMBER: US 60/005,529

; PRIOR FILING DATE: 1995-10-18

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1984

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (374)...(1894)

US-08-733-622C-1

Query Match

Best Local Similarity 1.0%; Score 17; DB 4; Length 1984;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1236 GCAGGAGTGGAAAGA 1252
Db 1438 GCAGGAGTGGAAAGA 1454

RESULT 88

US-09-976-594-667

; Sequence 667, Application US/09976594

; Patent No. 6673549

; GENERAL INFORMATION:

; APPLICANT: Furness, Michael

; APPLICANT: Buchbinder, Jenny

; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

; FILE REFERENCE: PA-0041 US

; CURRENT APPLICATION NUMBER: US/09/976,594

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/240,409

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 1143

; SOFTWARE: PERL Program

; SEQ ID NO 667

; LENGTH: 2015

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6673549 240518.42

; NAME/KEY: unsure

; LOCATION: 748

; OTHER INFORMATION: a, t, c, g, or other

US-09-976-594-667

Query Match

Best Local Similarity 1.0%; Score 17; DB 4; Length 2015;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 591 TGAAGTCTGAGGAGC 607
Db 1526 TGAAGTCTGAGGAGC 1542

RESULT 89

US-08-948-705-5

; Sequence 5, Application US/08948705A

; Patent No. 6043084

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ASSOCIATED WITH COLON CANCER AND METHODS FOR DIAGNOSING AND

; FILE REFERENCE: LUD-5506.1-JEL/NDH

; CURRENT APPLICATION NUMBER: US/08/948,705A

; CURRENT FILING DATE: 1997-10-10

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 2162

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-08-948-705-5

Query Match

Best Local Similarity 1.0%; Score 17; DB 3; Length 2162;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 736 CGTGAGCTGCAGCGCA 752
Db 1024 CGTGAGCTGCAGCGCA 1040

RESULT 90

US-09-510-543-5

; Sequence 5, Application US/09510543

; Patent No. 6517837

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH COLON CANCER AND

; FILE REFERENCE: LUD-5506.1-JEL/NDH

; CURRENT APPLICATION NUMBER: US/09/510,543

; CURRENT FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 09/102,322

; PRIOR FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 22

; SEQ ID NO 5

; LENGTH: 2162

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-510-543-5

Query Match

Best Local Similarity 1.0%; Score 17; DB 4; Length 2162;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 736 CGTGAGCTGCAGCGCA 752
Db 1024 CGTGAGCTGCAGCGCA 1040

RESULT 91

US-09-620-312D-282

; Sequence 282, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinchaast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 282_ Length: 2230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)..(1152)
US-09-620-312D-282

Query Match 1.0%; Score 17; DB 4; Length 2230;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 492 AGCCAAAGGTTTGGGTA 508
Db 916 AGCCAAAGGTTTGGGTA 932

RESULT 92
US-08-948-705-4
; Sequence 4, Application US/08948705A
; Patent No. 6043084
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ASSOCIATED WITH COLON CANCER AND METHODS FOR DIAGNOSING AND
; FILE REFERENCE: LUD-5506-JEL/NDH
; CURRENT APPLICATION NUMBER: US/08/948,705A
; CURRENT FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 2236
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-08-948-705-4

Query Match 1.0%; Score 17; DB 3; Length 2236;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 736 CGTGAGCTGCAGCGGCA 752
Db 1024 CGTGAGCTGCAGCGGCA 1040
RESULT 93
US-09-510-543-4
; Sequence 4, Application US/09510543
; Patent No. 6517837
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH COLON CANCER AND
; FILE REFERENCE: LUD-5506.1-JEL/NDH
; CURRENT APPLICATION NUMBER: US/09/510,543
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/102,322
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 4
; LENGTH: 2236
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-510-543-4

Query Match 1.0%; Score 17; DB 4; Length 2236;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 736 CGTGAGCTGCAGCGGCA 752
Db 1024 CGTGAGCTGCAGCGGCA 1040

RESULT 94
US-08-948-705-8
; Sequence 8, Application US/08948705A
; Patent No. 6043084
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ASSOCIATED WITH COLON CANCER AND METHODS FOR DIAGNOSING AND
; FILE REFERENCE: LUD-5506-JEL/NDH
; CURRENT APPLICATION NUMBER: US/08/948,705A
; CURRENT FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-08-948-705-8

Query Match 1.0%; Score 17; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 736 CGTGAGCTGCAGCGGCA 752
Db 1024 CGTGAGCTGCAGCGGCA 1040

RESULT 95

US-09-510-543-8
; Sequence 8, Application US/09510543
; Patent No. 6517837

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH COLON CANCER AND
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING COLON CANCER

; FILE REFERENCE: LUD-5506.1-JEL/NDH

; CURRENT APPLICATION NUMBER: US/09/510,543

; PRIOR FILING DATE: 2000-02-22

; PRIOR FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 22

; SEQ ID NO 8

; LENGTH: 2289

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-510-543-8

Query Match 1.0%; Score 17; DB 4; Length 2289;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 CGTGAGCTCGACGGCA 752

|||||

Db 1024 CGTGAGCTCGACGGCA 1040

RESULT 96

US-09-510-543-9

; Sequence 9, Application US/09510543

; Patent No. 6517837

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH COLON CANCER AND
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING COLON CANCER

; FILE REFERENCE: LUD-5506.1-JEL/NDH

; CURRENT APPLICATION NUMBER: US/09/510,543

; PRIOR FILING DATE: 2000-02-22

; PRIOR FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 22

; SEQ ID NO 9

; LENGTH: 2409

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-510-543-9

Query Match

Best Local Similarity 100.0%; Score 17; DB 4; Length 2409;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 CGTGAGCTCGACGGCA 752

|||||

Db 1024 CGTGAGCTCGACGGCA 1040

RESULT 97

US-09-004-502-4

; Sequence 4, Application US/09004502

; Patent No. 5962263

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Shah, Purvi

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Dr.

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ For Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/004,502

; FILING DATE: Filed Herewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0456 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2434 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BRSTNOT01

; CLONE: 412453

US-09-004-502-4

Query Match

Best Local Similarity 100.0%; Score 17; DB 2; Length 2434;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1022 TCCAGCAGGCCAGGAG 1038

|||||

Db 983 TCCAGCAGGCCAGGAG 999

RESULT 98

US-09-360-125-4

; Sequence 4, Application US/09360125

; Patent No. 6235715

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Shah, Purvi

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Dr.

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ For Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/360,125

; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA: US/09/004,502
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0456 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2434 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT01
CLONE: 412453
US-09-360-125-4

Query Match 1.0%; Score 17; DB 3; Length 2434;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1022 TCCAGCAGGCCCGAGGAG 1038
DB 983 TCCAGCAGGCCCGAGGAG 999

RESULT 99
US-09-227-357-79
Sequence 79, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 2455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (2277)
OTHER INFORMATION: n equals a.t.g. or c
US-09-227-357-79

Query Match 1.0%; Score 17; DB 4; Length 2455;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1022 TCCAGCAGGCCCGAGGAG 1038
DB 991 TCCAGCAGGCCCGAGGAG 1007

RESULT 100
US-08-669-286-6
Sequence 6, Application US/08669286
Patent No. 6130060
GENERAL INFORMATION:
APPLICANT: NAKAMURA, SEIJI
APPLICANT: SAKURAI, TAKASHI
APPLICANT: NEZU, JUNI-ICHI
TITLE OF INVENTION: GENE ENCODING ADSEVERIN
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,286
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 230-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2630 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 79..2223
US-08-669-286-6

Query Match 1.0%; Score 17; DB 3; Length 2630;
Best Local Similarity 100.0%; Pred.No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 ACGGGGCTGCCAACA 39
Db 938 ACGGGGCTGCCAACA 954

Search completed: June 9, 2004, 00:56:48
Job time : 120.148 secs

This Page Blank (uspto)

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	1638	100.0	1638	5	AAH26501 Human low
2	1536	93.8	2523	4	ABA09008 Human LDL
3	1536	93.8	2523	4	AAK52874 Human pol
4	1536	93.8	5085	7	ACC46153 Human dit
5	1536	93.8	6577	7	ACC46130 Human dit
6	1439	87.9	4697	5	AAH26496 Human low
7	1286	78.5	4697	2	AAV32839 Human low
8	1042	63.6	2188	4	AAK51890 Human low
9	1028	62.8	1793	7	AAK234807 Coding se
10	890	54.3	1793	3	AAK352118 Human ade
11	890	54.3	1793	3	AAK35217 Human ade
12	890	54.3	1793	3	AAK21340 Human low
13	890	54.3	1793	3	AAK21339 Human low
14	890	54.3	1793	3	AAK297034 Human nuc
15	890	54.3	1793	7	AAK297033 Human nuc
16	723	44.1	1854	2	AAQ66987 Interleuk
17	686	41.9	5284	5	AAK65026 DNA clone
18	413	25.2	529	2	AAK88391 EST clone
19	336	20.5	22255	5	AAH26437 Human low
20	336	20.5	115756	7	ACD13448 Human DNA
21	140	8.5	405	4	AAK35582 Human car
22	140	8.5	405	9	ADE45661 Human car
23	59	3.6	4722	2	AAV32836 Rabbit lo

c 97 19 1.2 627 6 ABK77260 Abk77260 Bacillus
 98 19 1.2 668 4 AAH01784 AAH01784 Klebsiell
 99 1.2 714 4 AAH01785 AAH01785 Klebsiell
 c 100 19 1.2 1016 2 AAT29401 AAT29401 Lipase co

ALIGNMENTS

RESULT 1

AAH26501 ID AAH26501 standard; DNA, 1638 BP.

XX AC AAH26501;

XX DT 12-NOV-2001 (first entry)

XX DE Human low density lipoprotein binding protein 3 (LBP-3) DNA.

XX KW Low density lipoprotein binding protein 3; LBP-3; LDL; human;

XX KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;

XX KW ds.

XX OS Homo sapiens.

XX PN WO200164874-A2.

XX PD 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US006356.

XX PR 02-MAR-2000; 2000US-00517849.

XX PR 14-JUL-2000; 2000US-00616289.

XX PA (BOST-) BOSTON HEART FOUND INC.

XX PI Lees AM, Lees RS, Law SW, Arjona AA;

XX DR WPI; 2001-565505/63.

XX DR P-PSDB; AAB82808.

XX PT New isolated low density lipoprotein binding polypeptide for treating,

XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX PS Claim 8; Fig 8A; 143pp; English.

XX CC The present sequence is that of the coding region of the human gene (see

XX CC also AAH26496) encoding novel human low density lipoprotein binding

XX CC protein 2 (LBP-3, see AAB82808). The gene was isolated from a genomic DNA

XX CC library using LBP-3 cDNA as probe. The LBP-3 protein predicted from the

XX CC present sequence differs from that predicted from a cDNA clone (see

XX CC AAB82804) in that it contains an additional 16 amino acids at its N-

XX CC terminus (the cDNA is a 5' truncation) and has Asn at amino acid position

XX CC 130 rather than Tyr. LBP-3 nucleic acids are among claimed

XX CC polynucleotides of the invention that encode novel polypeptides capable

XX CC of binding to native and methylated LDL. Also claimed are isolated LBP

XX CC polypeptides, and biologically active fragments and analogues of them, as

XX CC well as expression vectors, cells and methods of producing the LBPs.

XX CC Methods of determining if an animal is at risk for atherosclerosis,

XX CC methods for evaluating an agent for use in treating atherosclerosis, and

XX CC methods for treating a cell having an abnormality in structure or

XX CC metabolism of LBP are claimed. Pharmaceutical compositions comprising an

XX CC LBP polypeptide or nucleic acid, and vaccine compositions, are also

XX CC claimed

XX SQ Sequence 1638 BP; 490 A; 417 C; 520 G; 211 T; 0 U; 0 Other;

Query Match 100.0%; Score 1638; DB 5; Length 1638;

Best Local Similarity 100.0%; Pred No. 0;

Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAACCAAGACAAAAGACGGGGCTGCCAAATCCAAATCCAAAGACGACCA 60

Db 1 ATGAGAACCAAGACAAAAGACGGGGCTGCCAAATCCAAATCCAAAGACGACCA 60
 QY 61 GGACAAACCGGAAGCAGGACCCGAGGAGCCAGAGCGGCCAGCCAGCGGGCTCTCGCA 120
 Db 61 GGACAAACCGGAAGCAGGACCCGAGGAGCCAGAGCGGCCAGCCAGCGGGCTCTCGCA 120
 QY 121 GTAGAGCAGAGAGTCCCGGCGAGCAGCAGCTCCTCGGAGCGCGGGGTCTCAAGCC 180
 Db 121 GTAGAGCAGAGAGTCCCGGCGAGCAGCAGCTCCTCGGAGCGCGGGGTCTCAAGCC 180
 QY 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGCCGCAACTGAA 240
 Db 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGCCGCAACTGAA 240
 QY 241 GACATCTGAGCAGACATCTGTGTGACAAATACAGGGGGGGGGGGGGGGGGGGGGGCA 300
 Db 241 GACATCTGAGCAGACATCTGTGTGACAAATACAGGGGGGGGGGGGGGGGGGGGGGCA 300
 QY 301 CAGGTGAGCGCGCTGAACCCGAGATCCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 360
 Db 301 CAGGTGAGCGCGCTGAACCCGAGATCCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 360
 QY 361 GGGGAGCTTGAACCACTCCAGTAGTCAATGGAGAGAGAAACCTCCAGGGGGGATCCA 420
 Db 361 GGGGAGCTTGAACCACTCCAGTAGTCAATGGAGAGAGAAACCTCCAGGGGGGATCCA 420
 QY 421 AACACAGAGAGATCCCGGCGAGGTGACGAGTCCGAGACCGAGACCATCGAAGGCCACAG 480
 Db 421 AACACAGAGAGATCCCGGCGAGGTGACGAGTCCGAGACCGAGACCATCGAAGGCCACAG 480
 QY 481 GAGAAGAAAAAGCAAGGGTTTGGGTAAGAGATCACGTTGTGATGACAGATTGAAT 540
 Db 481 GAGAAGAAAAAGCAAGGGTTTGGGTAAGAGATCACGTTGTGATGACAGATTGAAT 540
 QY 541 ACTCTGAGTACCCAGAGGAGAGAGCTGGCTGCTGTGCAAGAGATGCTGAACCTGCTG 600
 Db 541 ACTCTGAGTACCCAGAGGAGAGAGCTGGCTGCTGTGCAAGAGATGCTGAACCTGCTG 600
 QY 601 GAGGAGCACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
 Db 601 GAGGAGCACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
 QY 661 GTGCAAGAGAGAGACCACTGCGCGGTGAGCAGACAGAGCGCTCTGCGGCCGAGCAAG 720
 Db 661 GTGCAAGAGAGAGACCACTGCGCGGTGAGCAGACAGAGCGCTCTGCGGCCGAGCAAG 720
 QY 721 CTTGAGAGCCTATCCGCTGAGCTCAGCGGCGACAAACCGCTCCCTCAAGGAAGAGTGTG 780
 Db 721 CTTGAGAGCCTATCCGCTGAGCTCAGCGGCGACAAACCGCTCCCTCAAGGAAGAGTGTG 780
 QY 781 CAGCGGCGCGGAGGAGGAGAGAGCGGAGAGAGTGAACCTCGACATTCAGAGGTGACA 840
 Db 781 CAGCGGCGCGGAGGAGGAGAGAGCGGAGAGAGTGAACCTCGACATTCAGAGGTGACA 840
 QY 841 CTGAATGACATTCAGCTGACAGTGGAAACAGCACAATGAGCGCAACTCCAAAGCTGCGCAA 900
 Db 841 CTGAATGACATTCAGCTGACAGTGGAAACAGCACAATGAGCGCAACTCCAAAGCTGCGCAA 900
 QY 901 GAGAACATGAGAGTGGCTGAGAGGCTCAAGAGCTGATGAGCAGATGATGAGTCCGGAG 960
 Db 901 GAGAACATGAGAGTGGCTGAGAGGCTCAAGAGCTGATGAGCAGATGATGAGTCCGGAG 960
 QY 961 GAGCATATCGAAGAGTCTTCAACACAGAGACCTACACAGCAGCTGGTGGATGCCAAG 1020
 Db 961 GAGCATATCGAAGAGTCTTCAACACAGAGACCTACACAGCAGCTGGTGGATGCCAAG 1020
 QY 1021 CTCAGCAGGCGCCAGGAGATGCTAAAGAGCGGAGAGCGGACCCAGCGGGGAGAGAT 1080
 Db 1021 CTCAGCAGGCGCCAGGAGATGCTAAAGAGCGGAGAGCGGACCCAGCGGGGAGAGAT 1080
 QY 1081 TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGCTGAGCTGATGAGCAGCAAGAG 1140
 Db 1081 TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGCTGAGCTGATGAGCAGCAAGAG 1140

QY 1141 ACCACCTGAAGCAACAGCTTCCCTTATACACAGAGAAGTTTGGAGTTCAGAACACA 1200
Db 1141 ACCACCTGAAGCAACAGCTTCCCTTATACACAGAGAAGTTTGGAGTTCAGAACACA 1200
QY 1201 CTTTCCAAAAGCAGCGAGTATTACCAATTCACAGAGATGGAAGATGACTAAG 1260
Db 1201 CTTTCCAAAAGCAGCGAGTATTACCAATTCACAGAGATGGAAGATGACTAAG 1260
QY 1261 AAGATCAAGAAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGTGGGAGAGCAGCAAC 1320
Db 1261 AAGATCAAGAAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGTGGGAGAGCAGCAAC 1320
QY 1321 AAGCCCTCTTTCAGATGCTGAGGAGAAACAGTCCGGGATAAGAACTGAGGCGCTG 1380
Db 1321 AAGCCCTCTTTCAGATGCTGAGGAGAAACAGTCCGGGATAAGAACTGAGGCGCTG 1380
QY 1381 CAGGTAAAATCCAAACGGCTGGAGAGAGCTGTGCGGGCACTGCAGACAGCGCAATGAC 1440
Db 1381 CAGGTAAAATCCAAACGGCTGGAGAGAGCTGTGCGGGCACTGCAGACAGCGCAATGAC 1440
QY 1441 CTGAACAAGAGGTACAGACCTGAGTGTGCTGGTGGCCAGGGTCCCTCTACAGAGTGGC 1500
Db 1441 CTGAACAAGAGGTACAGACCTGAGTGTGCTGGTGGCCAGGGTCCCTCTACAGAGTGGC 1500
QY 1501 CCTGAGAGAGGCGCAGAGGCGCTGGGCTCAAGACCCAGTCCCGGAGGTTCACAGAA 1560
Db 1501 CCTGAGAGAGGCGCAGAGGCGCTGGGCTCAAGACCCAGTCCCGGAGGTTCACAGAA 1560
QY 1561 GCGCTTGTGTACCCAGGAGCAGCGACAGAGCATAGGCGAGTGGGCTCAAGAG 1620
Db 1561 GCGCTTGTGTACCCAGGAGCAGCGACAGAGCATAGGCGAGTGGGCTCAAGAG 1620
QY 1621 CCCACCTCGCAGGGCC 1638
Db 1621 CCCACCTCGCAGGGCC 1638

RESULT 2

ABA09008

ID ABA09008 standard; cDNA; 2523 BP.

XX ABA09008;

XX 11-JAN-2002 (first entry)

XX Human LDL binding protein homologue-encoding cDNA, SEQ ID NO:784.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytotaxtic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.

OS Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003800.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX

XX

XX

XX

XX

XX

XX

XX

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PSDB; ABB11764.

Human proteins and DNA encoding sequences useful for preventing, treating
or ameliorating a medical condition in a mammalian subject e.g. arthritis
and cancer.

Claim 1; Page 693-694; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides,
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;

haematopoiesis regulatory activity; tissue growth factor activity;
immunomodulatory activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a cDNA encoding a
novel human polypeptide of the invention

Sequence 2523 BP; 666 A; 626 C; 794 G; 437 T; 0 U; 0 Other;

Query Match 93.8%; Score 1536; DB 4; Length 2523;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1636; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGAACCAAGACAAAAGACGCGGCTCCCAACCAATCCAAAAGACAGCCCA 60

Db 119 ATGAAGAACCAAGACAAAAGACGCGGCTCCCAACCAATCCAAAAGACAGCCCA 178

QY 61 GGACAAACCGAAGCAGGACCCGAGGAGCCAGGAGCCAGGCGGCTCCTTGCA 120

Db 179 GGACAAACCGAAGCAGGACCCGAGGAGCCAGGAGCCAGGCGGCTCCTTGCA 238

QY 121 GTAGAAGCAGAAAGTCCCGGAGCAGCCAGGCTCTCTCGAAGCCGAGGCTCAAGCC 180

Db 239 GTAGAAGCAGAAAGTCCCGGAGCAGCCAGGCTCTCTCGAAGCCGAGGCTCAAGCC 298

QY 181 AGAACGGCTCAGTCTGGGGCCCTTCGTATCTCTGAGGAGCTGAGCCGCAACTGAA 240

Db 299 AGAACGGCTCAGTCTGGGGCCCTTCGTATCTCTGAGGAGCTGAGCCGCAACTGAA 358

241 GACATCTGAGCATACTCTGTGGAATTAACACAGGGGGGGCCCCCGGCGAGGATGGGGCA 300
359 GACATCTGAGCATACTCTGTGGAATTAACACAGGGGGGGCCCCCGGCGAGGATGGGGCA 418
301 CAGGGTGAAGCCGCTGAACCCGGAAGATGACAGAAAGTCCCGGACCTATGTGGCAAGGAAT 360
419 CAGGGTGAAGCCGCTGAACCCGGAAGATGACAGAAAGTCCCGGACCTATGTGGCAAGGAAT 478
361 GGGGAGCTTGAACCACTCCAGTAGTCAATGGAGAGAGAACTCCCAAGGGGGATCCA 420
479 GGGGAGCTTGAACCACTCCAGTAGTCAATGGAGAGAGAACTCCCAAGGGGGATCCA 538
421 AACACAGAGAGATCCGGCAGAGTGAAGGTCGGAGACCGAGACCATCGAAGGCCACAG 480
539 AACACAGAGAGATCCGGCAGAGTGAAGGTCGGAGACCGAGACCATCGAAGGCCACAG 598
481 GAGAAGAAAAAAGCCAAAGGGTTTGGGTAAAGAGATCACGTTGCTGATGCAGACATTGAAT 540
599 GAGAAGAAAAAAGCCAAAGGGTTTGGGAGAGAGATCACGTTGCTGATGCAGACATTGAAT 658
541 ACTCTGAGTACCCAGAGAGAGAGTGGCTGCTCTGTGCAAGAGTATGCTGAATCTGCTG 600
659 ACTCTGAGTACCCAGAGAGAGAGTGGCTGCTCTGTGCAAGAGTATGCTGAATCTGCTG 718
601 GAGGAGCAGCCGAATTCACAGAGAGAGATGAAGCTCTTACAGAAAAAGCAGAGCCAGCTG 660
719 GAGGAGCAGCCGAATTCACAGAGAGAGATGAAGCTCTTACAGAAAAAGCAGAGCCAGCTG 778
661 GTGCAAGAGAGAGACCACTTGCCTGAGTGAACAGCAAGCCGCTCTGCCCCGCGAGCAAG 720
779 GTGCAAGAGAGAGACCACTTGCCTGAGTGAACAGCAAGCCGCTCTGCCCCGCGAGCAAG 838
721 CTTGAGAGCTTATGCTGAGTGAACAGCAAGCCGCTCTTCAAGAGAGAGTGTG 780
839 CTTGAGAGCTTATGCTGAGTGAACAGCAAGCCGCTCTTCAAGAGAGAGTGTG 898
781 CAGCGGGCCCCGGGAGGAG 840
899 CAGCGGGCCCCGGGAGGAG 958
841 CTGAATGACATTCAGCTGAGTGAACAGCAAGCCGCTCTTCAAGAGAGAGTGTG 900
959 CTGAATGACATTCAGCTGAGTGAACAGCAAGCCGCTCTTCAAGAGAGAGTGTG 1018
901 GAGAATGAGAGCTGCTGAGAGCTCAAGAGCTGATGAGCAGTATGAGCTGCGCGAG 960
1019 GAGAATGAGAGCTGCTGAGAGCTCAAGAGCTGATGAGCAGTATGAGCTGCGCGAG 1078
961 GAGCATATGACAAAGTCTTCAACACAGAGACCTTCAACAGAGCTGTGTGATGCCAAG 1020
1079 GAGCATATGACAAAGTCTTCAACACAGAGACCTTCAACAGAGCTGTGTGATGCCAAG 1138
1021 CTCAGCAGCCCGAGAGAGATGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080
1139 CTCAGCAGCCCGAGAGATGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1198
1081 TTTCTCTGAAAGAGCAGTAGAGTCCCAAGAGATGCTGAGCTGATGAAGCAGCAAGAG 1140
1199 TTTCTCTGAAAGAGCAGTAGAGTCCCAAGAGATGCTGAGCTGATGAAGCAGCAAGAG 1258
1141 ACCACCTGAGCAACAGCTTGCCTTATACAGAGAGAGTTCAGAGAGTTCAGAGACACA 1200
1259 ACCACCTGAGCAACAGCTTGCCTTATACAGAGAGAGTTCAGAGAGTTCAGAGACACA 1318
1201 CTTTCCAAAGAGCAGGAGTATTCACCATTCAGCAGAGAGATGGAAGAGTGAAGTAAAG 1260
1319 CTTTCCAAAGAGCAGGAGTATTCACCATTCAGCAGAGAGATGGAAGAGTGAAGTAAAG 1378
1261 AAGATCAAGAGAGTGAAG 1320
1379 AAGATCAAGAGAGTGAAG 1438
1321 AAGGCCCTGTTGAGATGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

1439 AAGGCCCTGTTGAGATGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1498
1381 CAGGTAAATAATCCAAACCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1499 CAGGTAAATAATCCAAACCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558
1441 CTGAACAG 1500
1559 CTGAACAG 1618
1501 CCTGAG 1560
1619 CCTGAG 1678
1561 GCGCTTGTCTACCCAG 1620
1679 GCGCTTGTCTACCCAG 1738
1621 CCCACCTCCGCGCAGGGCC 1638
1739 CCCACCTCCGCGCAGGGCC 1756

RESULT 3

AAK52874

ID AAK52874 standard; cDNA; 2523 BP.

XX

AC AAK52874;

XX

DT 06-NOV-2001 (first entry)

XX

Human polynucleotide SEQ ID NO 2403.

XX

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
tissue growth factor; immunomodulatory; cancer; leukaemia;
nervous system disorder; arthritis; inflammation; ss.

XX

Homo sapiens.

XX

WO200157190-A2.

XX

09-AUG-2001.

XX

05-FEB-2001; 2001WO-US004098.

XX

03-FEB-2000; 2000US-00496914.

XX

27-APR-2000; 2000US-00560875.

XX

19-JUL-2000; 2000US-00598075.

XX

01-SEP-2000; 2000US-00620325.

XX

15-SEP-2000; 2000US-00654936.

XX

20-OCT-2000; 2000US-00693325.

XX

30-NOV-2000; 2000US-00728422.

XX

(HYSB-) HYSB INC.

XX

Tang Y, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX

Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang Z;

XX

Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

WPI; 2001-476283/51.

XX

P-FSDB; AAM79741.

XX

Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX

in diagnosis and gene therapy.

XX

Claim 1; Page 4687-4688; 6221pp; English.

XX

The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX

encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to

XX

cytokine, cell proliferation or cell differentiation or which may induce

XX

CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX

Seq Sequence 2523 BP; 666 A; 626 C; 794 G; 437 T; 0 U; 0 Other;

Query Match 93.8%; Score 1536; DB 4; Length 2523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAAGAACCAAGACAAAGAACGCGGTGCGCAACCAATCCAAATCCAAAGACGACCA 60
DB 119 ATGAAGAACCAAGACAAAGAACGCGGTGCGCAACCAATCCAAATCCAAAGACGACCA 178
QY 61 GGACAAACCGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCCAGGCGGCTCTGCA 120
DB 179 GACAAACCGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCCAGGCGGCTCTGCA 238
QY 121 GTAGAAGCAGAAAGTCCCGGAGCAGCAGGCTCTCGGAGCCCGAGGCGGTCTCAAGCC 180
DB 239 GTAGAAGCAGAAAGTCCCGGAGCAGCAGGCTCTCGGAGCCCGAGGCGGTCTCAAGCC 298
QY 181 AGAAGCGCTCAGTCTGGGGCCCTTCTGTATGCTCTGAGGAGCTCAGCGCCCACTGGAA 240
DB 299 AGAAGCGCTCAGTCTGGGGCCCTTCTGTATGCTCTGAGGAGCTCAGCGCCCACTGGAA 358
QY 241 GACATCTGAGCACATATCTGTGACAAATACAGGGGGGGCCCGGAGAGATGGGGCA 300
DB 359 GACATCTGAGCACATATCTGTGACAAATACAGGGGGGGCCCGGAGAGATGGGGCA 418
QY 301 CAGGTGACCGCGTGAACCCGAGATGAGAGAGTCCCGACCTATGTGCGAAGGAT 360
DB 419 CAGGTGACCGCGTGAACCCGAGATGAGAGAGTCCCGACCTATGTGCGAAGGAT 478
QY 361 GGGGAGCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCCCTCCAAAGGGGATCCA 420
DB 479 GGGGAGCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCCCTCCAAAGGGGATCCA 538
QY 421 AACACAGAGAGATCCGGCAGAGTGAAGAGTCCGAGAGCCGAGACATCGAAGGCCACAG 480
DB 539 AACACAGAGAGATCCGGCAGAGTGAAGAGTCCGAGAGCCGAGACATCGAAGGCCACAG 598
QY 481 GAGAGAAAAGCCCAAGGTTTGGTAAAGGATCAGCTTCTGATGACAGATTTGAAT 540
DB 599 GAGAGAAAAGCCCAAGGTTTGGTAAAGGATCAGCTTCTGATGACAGATTTGAAT 598
QY 541 ACTCTGAGTACCCAGAGGAGAGTGGTGTCTGTGCAAGAGTATGCTGAACTGTGTG 600
DB 659 ACTCTGAGTACCCAGAGGAGAGTGGTGTCTGTGCAAGAGTATGCTGAACTGTGTG 718
QY 601 GAGGACACCGGATTCAGAGAGAGATGAGCTCTACAGAAAGCAGAGCCAGCTG 660
DB 719 GAGGACACCGGATTCAGAGAGAGATGAGCTCTACAGAAAGCAGAGCCAGCTG 778
QY 661 GTCAAGAGAGAGACCACTTGGCGGTGAGCAGCAGCAAGCGGCTCTGCGCCGCGCAGAG 720
DB 779 GTCAAGAGAGAGACCACTTGGCGGTGAGCAGCAGCAAGCGGCTCTGCGCCGCGCAGAG 838
QY 721 CTTGAGAGCTATGCGGTGAGCTGAGCGGCACACCGCTCTCTCAGGAGAGAGGTGTG 780
DB 839 CTTGAGAGCTATGCGGTGAGCTGAGCGGCACACCGCTCTCTCAGGAGAGAGGTGTG 898
QY 781 CAGCGGCGCGGAGGAGGAGAGAGCGCAGAGGAGTCACTCGACATTCAGAGGTGACA 840
DB 899 CAGCGGCGCGGAGGAGGAGAGAGCGCAGAGGAGTCACTCGACATTCAGAGGTGACA 958
```

```
QY 841 CTGAATGACANTCAGCTGCAGATGGAACAGCAATGAGGGCACTCCAAGCTGCGCCAA 900
DB 959 CTGAATGACANTCAGCTGCAGATGGAACAGCAATGAGGGCACTCCAAGCTGCGCCAA 1018
QY 901 GAGAATCATGAGCTGGCTGAGAGGCTCAAGAACTGATTGAGCAGTATGAGCTGCGCGAG 960
DB 1019 GAGAATCATGAGCTGGCTGAGAGGCTCAAGAACTGATTGAGCAGTATGAGCTGCGCGAG 1078
QY 961 GAGCATATCGACAAAGTCTTCAAAACAAGGACCTCAACAGCAGCTGCTGGTGGATGCAAG 1020
DB 1079 GAGCATATCGACAAAGTCTTCAAAACAAGGACCTCAACAGCAGCTGCTGGTGGATGCAAG 1138
QY 1021 CTCAGAGAGAGAGAGAGATGCTTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAT 1080
DB 1139 CTCAGAGAGAGAGAGAGATGCTTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAT 1198
QY 1081 TTTCTCTCTGAAAGAGGAGGAGTAGAGTCCAGAGAGATGTGTGAGCTGATGAAGCAGCAAG 1140
DB 1199 TTTCTCTCTGAAAGAGGAGGAGTAGAGTCCAGAGAGATGTGTGAGCTGATGAAGCAGCAAG 1258
QY 1141 ACCCACCTGAAAGCAACAGCTTGGCTTATACAGAGAGAGTGTGAGGAGTTCAGAGACACA 1200
DB 1259 ACCCACCTGAAAGCAACAGCTTGGCTTATACAGAGAGAGTGTGAGGAGTTCAGAGACACA 1318
QY 1201 CTTTCCAAAAGAGCAGGAGTATTACCAATTCAGCAGAGAGATGGAAGAGATGACTAAG 1260
DB 1319 CTTTCCAAAAGAGCAGGAGTATTACCAATTCAGCAGAGAGATGGAAGAGATGACTAAG 1378
QY 1261 AAGATCAAGAGAGTGGAGAAAGAAACCAATGTACCGTCCCGGTGGGAGAGAGCAAC 1320
DB 1379 AAGATCAAGAGAGTGGAGAAAGAAACCAATGTACCGTCCCGGTGGGAGAGAGCAAC 1438
QY 1321 AAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATGAAGAACTGGAGGGCTG 1380
DB 1439 AAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATGAAGAACTGGAGGGCTG 1498
QY 1381 CAGGTAAAAATCCAAACCGCTGGAGAGCTGTCCCGGCACTGCAGACAGAGCGCAATGAC 1440
DB 1499 CAGGTAAAAATCCAAACCGCTGGAGAGCTGTCCCGGCACTGCAGACAGAGCGCAATGAC 1558
QY 1441 CTGAACAGAGAGTACAGGAGCTGAGTGTGTGGCCAGAGGCTCCCTCACTGACAGTGGC 1500
DB 1559 CTGAACAGAGAGTACAGGAGCTGAGTGTGTGGCCAGAGGCTCCCTCACTGACAGTGGC 1618
QY 1501 CCTCAGAGAGAGGAGGAGGAGGCTGAGGCTCAAGACACCGAGCTCCCGGCTCAGAGAA 1560
DB 1619 CCTCAGAGAGAGGAGGAGGAGGCTGAGGCTCAAGACACCGAGCTCCCGGCTCAGAGAA 1678
QY 1561 GCGCTTGTCTACCCAGAGGAGCAGCAGCAGCAGAGCAGCAGATCAGGCGCAGATGGGCTCAAGAG 1620
DB 1679 GCGCTTGTCTACCCAGAGGAGCAGCAGCAGCAGCAGCAGATCAGGCGCAGATGGGCTCAAGAG 1738
QY 1621 CCCACCTCCCGCAGGGCC 1638
DB 1739 CCCACCTCCCGCAGGGCC 1756
```

RESULT 4

ACC46153

ID ACC46153 standard; cDNA; 5085 BP.

XX ACC46153;

XX 02-JUN-2003 (first entry)

XX Human dithp extracellular signalling protein-encoding cDNA.

DE Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;

XX cancer; cell proliferative disorder; autoimmune disorder;

KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;

KW neurological disorder; gastrointestinal disorder; transport disorder;

KW connective tissue disorder; drug screening; proteome analysis;

KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;

QY 1081 TTTCTCTGAAGAGCGGCTAGAGTCCAGAGGATGTGTGAGCTGATGAGCAGCAGAG 1140
Db 1360 TTTCTCTGAAGAGCGGCTAGAGTCCAGAGGATGTGTGAGCTGATGAGCAGCAGAG 1419
QY 1141 ACCCACTTGAAGCAACAGCTTCCCTATACAGAGAAAGTTTGGAGCTTCCAGAACACA 1200
Db 1420 ACCCACTTGAAGCAACAGCTTCCCTATACAGAGAAAGTTTGGAGCTTCCAGAACACA 1479
QY 1201 CTTTCCAAAGAGCAGGAGTATTCACCATTCAGAGAGGAGATGGAAGATGACTTAAG 1260
Db 1480 CTTTCCAAAGAGCAGGAGTATTCACCATTCAGAGAGGAGATGGAAGATGACTTAAG 1539
QY 1261 AAGATCAAGAGCTGGAGAGAAACCCATGTACCGTCCCGTGGAGAGCAGCAAC 1320
Db 1540 AAGATCAAGAGCTGGAGAGAAACCCATGTACCGTCCCGTGGAGAGCAGCAAC 1599
QY 1321 AAGGCCCTGCTTGAAGTGGCTGAGAGAGAAACAGTCCCGGATTAAGAACTGAGAGGCTG 1380
Db 1600 AAGGCCCTGCTTGAAGTGGCTGAGAGAGAAACAGTCCCGGATTAAGAACTGAGAGGCTG 1659
QY 1381 CAGGTAAATCCAAACGGCTGAGAGAGCTGTCGGGGCACTGCACAGAGGCAATCAC 1440
Db 1660 CAGGTAAATCCAAACGGCTGAGAGAGCTGTCGGGGCACTGCACAGAGGCGCAATGAC 1719
QY 1441 CTGAACAGAGGGGTACAGGACCTGAGTGTGTGTGCTGAGGCTCCCTCACTGACAGTGGC 1500
Db 1720 CTGAACAGAGGGGTACAGGACCTGAGTGTGTGTGCTGAGGCTCCCTCACTGACAGTGGC 1779
QY 1501 CTGAGAGAGGCGCAGAGGGGCTGGGCTCAAGACCCAGTCCCGCAGGCTCAGAA 1560
Db 1780 CTGAGAGAGGCGCAGAGGGGCTGGGCTCAAGACCCAGTCCCGCAGGCTCAGAA 1839
QY 1561 GCGCTTGTCTCCAGAGGACCGAGCAGAGAGATCAGGCGAGTCCGGGCTCAAGAG 1620
Db 1840 GCGCTTGTCTCCAGAGGACCGAGCAGAGAGATCAGGCGAGTCCGGGCTCAAGAG 1899
QY 1621 CCCACTCCCGCAGGGCC 1638
Db 1900 CCCACTCCCGCAGGGCC 1917

RESULT 5
ACC46130
ID ACC46130 standard; cDNA; 6577 BP.
XX AC ACC46130;
XX
XX
DT 02-JUN-2003 (first entry)
DE Human dithp extracellular signalling protein-encoding cDNA.
XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW extracellular signalling; gene; ss.
XX Homo sapiens.
XX
XX WO200297031-A2.
XX
XX
PD 05-DEC-2002.
XX
XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR

PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
XX Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
XX Peralta CH, David MH, Lewis SA, Chen AU, Panzer SR, Harris B;
XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
XX P-PSDB; ABR41186.
XX
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
XX test compound which specifically binds to a polypeptide encoded by human
XX diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
XX Claim 2; SEQ ID NO 51; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
XX polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
XX proteins (Dithp; ABR41136-ABR41812). The invention also relates to
XX polynucleotide sequences at least 90% identical to the dithp cDNA
XX sequences of the invention; recombinant vectors, host cells and
XX transgenic organisms comprising a dithp nucleic acid sequence; the
XX recombinant production of Dithp proteins; antibodies specific for Dithp
XX proteins; microarrays comprising dithp nucleic acid sequences; methods of
XX detecting dithp nucleotide and protein sequences; methods of screening
XX for compounds which specifically bind a Dithp protein; and methods of
XX assessing the toxicity of test compounds using a dithp hybridisation
XX probe. Dithp nucleic acid sequences and Dithp proteins may be used in the
XX diagnosis of a wide variety of conditions including cancer and other cell
XX proliferative disorders; autoimmune or inflammatory disorders; bacterial,
XX viral, fungal or parasitic infections; hormonal disorders; metabolic
XX disorders; neurological disorders; gastrointestinal disorders; transport
XX disorders; and connective tissue disorders. They may also be used to
XX screen for modulators of protein activity or gene expression. Dithp
XX proteins can additionally be used in analysis of the proteome of a tissue
XX or cell type and to induce antibodies. The dithp nucleic acids are
XX additionally useful in somatic or germline gene therapy of the disorders
XX mentioned above, as a source of antisense sequences, as a source of
XX probes and primers, in genotyping and identification of individuals, in
XX the generation of transgenic animal models of human disease or knock in
XX humanised animals, in toxicological testing, and in transcript imaging.
XX The present sequence represents a dithp cDNA encoding a Dithp protein
XX which has extracellular signalling activity. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6577 BP; 1496 A; 1701 C; 1836 G; 1540 T; 0 U; 4 Other;
XX
XX Query Match 93.8%; Score 1536; DB 7; Length 6577;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1836; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAAGAACCAAGACAAAGAACGCGGCTGCCAAACATCCATCCAAAGAGAGCCCA 60
Db 280 ATGAAGAACCAAGACAAAGAACGCGGCTGCCAAACATCCATCCAAAGAGAGCCCA 339
QY 61 GGACAAACCGGAGCAGGACCCGAGGAGCCAGGAGCGGCCAGGAGGCTCTTGA 120
Db 340 GGACAAACCGGAGCAGGACCCGAGGAGCCAGGAGCGGCCAGGAGGCTCTTGA 399
QY 121 GTAGAAGCAGAGAGTCCCGGCGAGCAGGAGCTCTCGAAGCCGAGGCTCAAGGCC 180
Db 400 GTAGAAGCAGAGAGTCCCGGCGAGCAGGAGCTCTCGAAGCCGAGGCTCAAGGCC 459
QY 181 AGACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCCGCCAAGTGA 240

QY	1547	CCAGGGTACAGAGCGCTTGTACCCAGGAGGACCCAGCAGCAGAGCATCAGGCCGAGA	1606
DB	1501	CCAGGGTACAGAGCGCTTGTACCCAGGAGGACCCAGCAGCAGAGCATCAGGCCGAGA	1560
QY	1607	CTGGGCTCAAGAGCGCCACCTCCGCCAGGGCC	1638
DB	1561	CTGGGCTCAAGAGCGCCACCTCCGCCAGGGCC	1592
RESULT 8			
AAK51890			
ID	AAK51890 standard; cDNA; 2188 BP.		
XX	AAK51890;		
AC	AAK51890;		
DT	06-NOV-2001 (first entry)		
XX	Human polynucleotide SEQ ID NO 435.		
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US004098.		
XX			
PR	03-FEB-2000; 2000US-00496914.		
PR	27-APR-2000; 2000US-00560875.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
XX			
PA	(HYSEQ-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;		
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;		
XX			
DR	WPI; 2001-476283/51.		
DR	P-PSDB; AAM78757.		
XX			
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful		
PT	in diagnosis and gene therapy.		
XX			
PS	Claim 1; Page 1605-1607; 6221pp; English.		
XX			
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111		
CC	(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the		
CC	sequence listing were missing at the time of publication		
XX			
SQ	Sequence 2188 BP; 541 A; 559 C; 669 G; 418 T; 0 U; 1 Other;		
Query Match 63.6%; Score 1042; DB 4; Length 2188;			
Best Local Similarity 100.0%; Pred. No. 0;			

Matches 1042; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	597	GCTGGAGGACACCGGAATTCACAGAGAGATGAAGCTCTCTACAGAAAAGCAGAGCCA	556
DB	381	GCTGGAGGACACCGGAATTCACAGAGAGATGAAGCTCTCTACAGAAAAGCAGAGCCA	440
QY	657	GCTGGTCAAGAGAGGACCACTGCGGTGAGCAGCAAGGCGCTCTGCGCCGACG	716
DB	441	GCTGGTCAAGAGAGGACCACTGCGGTGAGCAGCAAGGCGCTCTGCGCCGACG	500
QY	717	CAAGCTTGAGAGCTATGCCGTGAGCTGACGCGCAACCCGCTCCCTCAAGGAAGG	776
DB	501	CAAGCTTGAGAGCTATGCCGTGAGCTGACGCGCAACCCGCTCCCTCAAGGAAGG	560
QY	777	TGTGACGGGGCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	836
DB	561	TGTGACGGGGCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	620
QY	837	GACACTCAATGACATTCAGCTGACATGGAACAGCAATGAGCGCAATCCCAAGTGG	896
DB	621	GACACTCAATGACATTCAGCTGACATGGAACAGCAATGAGCGCAATCCCAAGTGG	680
QY	897	CCAAGAAACATGAGCTGCGCTGAGAGCTCAAGAGCTGATTGAGCAGTATGAGTGG	956
DB	681	CCAAGAAACATGAGCTGCGCTGAGAGCTCAAGAGCTGATTGAGCAGTATGAGTGG	740
QY	957	CGAGGACATATCGACAAAGTCTTCAAAACAAGGACCTACACAGCAGCTGTTGATGC	1016
DB	741	CGAGGACATATCGACAAAGTCTTCAAAACAAGGACCTACACAGCAGCTGTTGATGC	800
QY	1017	CAAGCTCCAGCAGCCCGAGGAGATGCTAAAGAGGAGGAGAGAGCGGCACCGGGGAA	1076
DB	801	CAAGCTCCAGCAGCCCGAGGAGATGCTAAAGAGGAGGAGAGAGCGGCACCGGGGAA	860
QY	1077	GGATTTCTCTGAAAGAGCAGTAGAGTCCAGAGATGTTGAGCTGATGAGCAGCA	1136
DB	861	GGATTTCTCTGAAAGAGCAGTAGAGTCCAGAGATGTTGAGCTGATGAGCAGCA	920
QY	1137	AGAGACCCACCTGAAGCAACAGCTTGCCCTATATACAGAGAAGTTTTCAGAGTTCCAGAA	1196
DB	921	AGAGACCCACCTGAAGCAACAGCTTGCCCTATATACAGAGAAGTTTTCAGAGTTCCAGAA	980
QY	1197	CACACTTCCAAAGAGCAGGAGTATTCACCAATTCACAGCAGGAGTGAAGAAGTAC	1256
DB	981	CACACTTCCAAAGAGCAGGAGTATTCACCAATTCACAGCAGGAGTGAAGAAGTAC	1040
QY	1257	TAAGAAGTCAAGAGCTGAGAGAGAAACACCATGTACCGGTCCCGGTGGGAGACAG	1316
DB	1041	TAAGAAGTCAAGAGCTGAGAGAGAAACACCATGTACCGGTCCCGGTGGGAGACAG	1100
QY	1317	CAACAAGGCCCTGCTTGCAGATGGCTGAGGAGAAACAGTCCGGGATAAAGACTGGAGGG	1376
DB	1101	CAACAAGGCCCTGCTTGCAGATGGCTGAGGAGAAACAGTCCGGGATAAAGACTGGAGGG	1160
QY	1377	CCTCAGGTAAATCCAGCGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCCCAA	1436
DB	1161	CCTCAGGTAAATCCAGCGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCCCAA	1220
QY	1437	TGACTCAACAAGAGGCTACAGGACCTGAGTGTGTTGGCGAGGGCTCCCTCACTGACAG	1496
DB	1221	TGACTCAACAAGAGGCTACAGGACCTGAGTGTGTTGGCGAGGGCTCCCTCACTGACAG	1280
QY	1497	TGGCCCTGAGAGGAGGCGCAGAGGGCTGGGGCTCAAGACCCAGCTCCCGCAGGGTAC	1556
DB	1281	TGGCCCTGAGAGGAGGCGCAGAGGGCTGGGGCTCAAGACCCAGCTCCCGCAGGGTAC	1340
QY	1557	AGAAGGCCCTTGTACCCAGGAGCAGCGAGCAGCAAGAGCATCAGGCCAGACTGGGCTCA	1616
DB	1341	AGAAGGCCCTTGTACCCAGGAGCAGCGAGCAGCAAGAGCATCAGGCCAGACTGGGCTCA	1400
QY	1617	AGAGCCCACTCCCGCAGGGCC	1638
DB	1401	AGAGCCCACTCCCGCAGGGCC	1422

[illegible]

DR WPI; 2000-205971/18.
XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension, or
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.
XX
PS Disclosure; Page 1259; 1343pp; English.
XX
CC The present invention describes a new composition comprising an antisense
CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have anti-inflammatory, antiallergic,
CC antisthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of the
CC ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA2313 to AAA3512 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC AAA33992) are specifically claimed ONs from the present invention. N.B.
CC Sequences given in the disclosure of the present invention do not match
CC up with their corresponding SEQ ID NO: sequences given in the sequence
CC listing
XX
SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;
Query Match 54.3%; Score 890; DB 3; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 749 GGCACAAACCCCTCCTTAAGAGAGAGAGTGTGAGCGGGCCCGGAGGAGGAGGAGAAC 808
DB 1656 GGCACAAACCCCTCCTTAAGAGAGAGAGTGTGAGCGGGCCCGGAGGAGGAGGAGAAC 1597
QY 809 GCAGGAGGTGACTCGCACTTCAGGTGACATGATGACATTTCAGCTGCAGATGGAAC 868
DB 1596 GCAGGAGGTGACTCGCACTTCAGGTGACATGATGACATTTCAGCTGCAGATGGAAC 1537
QY 869 AGCACAATAGCGCAACTCCAAAGCTGGCCAGAGAACATGAGCTGGCTGAGAGCTCA 928
DB 1536 AGCACAATAGCGCAACTCCAAAGCTGGCCAGAGAACATGAGCTGGCTGAGAGCTCA 1477
QY 929 AGRAGCTGATGAGCAGTATGAGCTGGCGGAGGAGCATATCGACAAAGTCTTCAAACACA 988
DB 1476 AGRAGCTGATGAGCAGTATGAGCTGGCGGAGGAGCATATCGACAAAGTCTTCAAACACA 1417
QY 989 AGGACCTTACACAGCAGCTGGTGGATGCCAAGCTCCAGAGGCCCGGAGGAGTCTAAAGG 1048
DB 1416 AGGACCTTACACAGCAGCTGGTGGATGCCAAGCTCCAGAGGCCCGGAGGAGTCTAAAGG 1357
QY 1049 AGGCAGAGAGCGGCACACGCGGAGNAGATTTCTCTGAAAGGCGGAGTGTAGTCTCC 1108
DB 1356 AGGCAGAGAGCGGCACACGCGGAGNAGATTTCTCTGAAAGGCGGAGTGTAGTCTCC 1297
QY 1109 AGAGGATGTGTGAGCTGATGAAGCAGCAGACAGACCCACCTGAAGCAACAGCTTGCCTAT 1168
DB 1296 AGAGGATGTGTGAGCTGATGAAGCAGCAGACAGACCCACCTGAAGCAACAGCTTGCCTAT 1237
QY 1169 ACACAGAGAGTTTGAGGAGTTCAGAACACACTTCCAAAGCAGGAGGTATTACCA 1228
DB 1236 ACACAGAGAGTTTGAGGAGTTCAGAACACACTTCCAAAGCAGGAGGTATTACCA 1177

QY 1229 CATTCAAGCAGGAGATGGAAGAGATGACTAAGAGATCAAGAGCTGGAGAAAACCA 1288
DB 1176 CATTCAAGCAGGAGATGGAAGAGATGACTAAGAGATCAAGAGCTGGAGAAAACCA 1117
QY 1289 CCATGTATCCCGTCCCGTGGGAGAGAGCAACAAAGCCCTGCTTGAGATGGCTGAGAGA 1348
DB 1116 CCATGTATCCCGTCCCGTGGGAGAGAGCAACAAAGCCCTGCTTGAGATGGCTGAGAGA 1057
QY 1349 AAACAGTCCCGGATTAAGAACTGGGCGCTGCAGGTAAATAATCCACGCTGGAGAGC 1408
DB 1056 AAACAGTCCCGGATTAAGAACTGGGCGCTGCAGGTAAATAATCCACGCTGGAGAGC 997
QY 1409 TGTGTCGGGCACTGTCACACAGAGCGCAATCACTGAACAGAGGTACAGGACTGAGTG 1468
DB 996 TGTGTCGGGCACTGTCACACAGAGCGCAATCACTGAACAGAGGTACAGGACTGAGTG 937
QY 1469 CTGTGTGCGAGGCTCCCTCACTGACAGTGGCCCTGAGAGGCGCCAGAGGGCCCTGGGG 1528
DB 936 CTGTGTGCGAGGCTCCCTCACTGACAGTGGCCCTGAGAGGCGCCAGAGGGCCCTGGGG 877
QY 1529 CTCAAGCACCAGCTCCCGCAGGGTCAACAGAGCGCTTCTACCCAGGAGCACCGAGCA 1588
DB 876 CTCAAGCACCAGCTCCCGCAGGGTCAACAGAGCGCTTCTACCCAGGAGCACCGAGCA 817
QY 1589 CAGAAGCATCAGGCGACAGCTGGGCTCAAGAGCCACCTCCGCCAGGGCC 1638
DB 816 CAGAAGCATCAGGCGACAGCTGGGCTCAAGAGCCACCTCCGCCAGGGCC 767
RESULT 12
AAP21340/c
ID AAP21340 standard; DNA; 1793 BP.
XX
AC AAP21340;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2907.
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; anti-inflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US008020.
XX
PR 06-APR-1999; 99US-0127958P.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX (NYCE/) NYCE J W.
XX
PI Nyce JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenosine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers
PT and respiratory obstructions.
XX
XX Disclosure; Page 1344; 1592pp; English.
XX

CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention
XX
XX
SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 54.3%; Score 890; DB 3; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 749 GGCACAAACCGCTCCCTCAAGGAAGAGGTGTCAGCGGGCCCGGAGGAGGAGGAGC 808
DB 1656 GGCACAAACCGCTCCCTCAAGGAAGAGGTGTCAGCGGGCCCGGAGGAGGAGGAGC 1597
QY 809 GCAAGAGGTGACCTCGACCTTCAGGTGACACTGAATGACATTGAGTGCAGATGGAAC 868
DB 1596 GCAAGAGGTGACCTCGACCTTCAGGTGACACTGAATGACATTGAGTGCAGATGGAAC 1537
QY 869 ASCACAATGAGCGCACTCCAAAGTCCGCAAGAGAGCATGAGTGGCTGAGAGGCTCA 928
DB 1536 AGCAATGAGCGCACTCCAAAGTCCGCAAGAGAGCATGAGTGGCTGAGAGGCTCA 1477
QY 929 AGAAGCTGATTGAGCAGTATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAACACA 988
DB 1476 AGAAGCTGATTGAGCAGTATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAACACA 1417
QY 989 AGGACCTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCCCGGAGGATGCTAAAGG 1048
DB 1416 AGGACCTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCCCGGAGGATGCTAAAGG 1357
QY 1049 AGGCAAGAGCGGCGACCGGAGAGAGGATTTCTCTCTGAAAGAGGAGTATGAGTCCC 1108
DB 1356 AGGCAAGAGCGGCGACCGGAGAGAGGATTTCTCTCTGAAAGAGGAGTATGAGTCCC 1297
QY 1109 AGAGATGTTGAGCTGATGAAGCAGCAGAGACCCACTGAAGCAACAGCTTGGCCCTAT 1168
DB 1296 AGAGATGTTGAGCTGATGAAGCAGCAGAGACCCACTGAAGCAACAGCTTGGCCCTAT 1237
QY 1169 ACACAGAGAGTTTGAAGGATTTCCAGACACACTTTCCAAAGAGCGAGGATTTACCA 1228
DB 1236 ACACAGAGAGTTTGAAGGATTTCCAGACACACTTTCCAAAGAGCGAGGATTTACCA 1177
QY 1229 CATTCAAGCAGAGATGGAAGAGTACTTAAGAGATCAAGAGCTGAGAGAGAAACCA 1288
DB 1176 CATTCAAGCAGAGATGGAAGAGTACTTAAGAGATCAAGAGCTGAGAGAGAAACCA 1117
QY 1289 CCATGTACCGGTCCCGGTGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGA 1348

DB 1116 CCATGTACCGGTCCCGGTGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGA 1057
QY 1349 AACAGTCCGGGATAAAGAACTGGAGGGCTGCGAGGTAAAAATCCACGCTGGAGAAGC 1408
DB 1056 AACAGTCCGGGATAAAGAACTGGAGGGCTGCGAGGTAAAAATCCACGCTGGAGAAGC 997
QY 1409 TGTGCCGGGCACTGCGACAGAGAGCGCAATCACTGAACAAGAGGGTACAGGACCTGAGTG 1468
DB 996 TGTGCCGGGCACTGCGACAGAGAGCGCAATCACTGAACAAGAGGGTACAGGACCTGAGTG 937
QY 1469 CTGTGTGCGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCGCCAGAGGGGCTGGGG 1528
DB 936 CTGTGTGCGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCGCCAGAGGGGCTGGGG 877
QY 1529 CTCAGACACCCAGCTCCCGCAGAGGTACAGAGCGGCTTCTCTACCCAGGAGCACCAGCA 1588
DB 876 CTCAGACACCCAGCTCCCGCAGAGGTACAGAGCGGCTTCTCTACCCAGGAGCACCAGCA 817
QY 1589 CAGAGATCAGCGCAGACTGGGCTCAAGAGCGCCACCTCCGCGAGGGCC 1638
DB 816 CAGAGATCAGCGCAGACTGGGCTCAAGAGCGCCACCTCCGCGAGGGCC 767

RESULT 13

AAF21339/C

ID AAF21339 standard; DNA; 1793 BP.

XX AAF21339;
XX

DT 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2906.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;
XX immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
XX respiratory obstruction; pulmonary obstruction; impeded respiration;
XX surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
XX respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
XX pulmonary hypertension; emphysema; pulmonary transplantation rejection;
XX chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
XX cancer; ss.

XX Homo sapiens.

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US008020.

XX 06-APR-1999; 99US-0127958P.

XX (UYEC-) UNIV EAST CAROLINA.
XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not trigger
XX adenosine receptors during metabolism, useful e.g. for treating cancers
XX and respiratory obstructions.

XX Disclosure; Page 1343; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
XX oligonucleotides and compositions (I) comprising them. In the antisense
XX oligonucleotides the A is replaced by a 'Universal' or alternative base.
XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
XX immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.

CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
 CC surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention
 XX

SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 54.3%; Score 890; DB 3; Length 1793;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	749	GGCACAAACCGCTCCCTCAAGGAAGAGGTGTGAGCGGCGCGGAGGAGGAGGAGC	808
DB	1656	GGCACAAACCGCTCCCTCAAGGAAGAGGTGTGAGCGGCGCGGAGGAGGAGGAGC	1597
QY	809	GCAAGAGGTGACTCGCACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAC	868
DB	1596	GCAAGAGGTGACTCGCACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAC	1537
QY	869	AGCACAAATGAGCGCACTTCCAGCTCGGCAAGAGAAATGAGCTGGCTGAGAGGCTCA	928
DB	1536	AGCACAAATGAGCGCACTTCCAGCTCGGCAAGAGAAATGAGCTGGCTGAGAGGCTCA	1477
QY	929	AGAAGCTGATTGAGCAGTATGAGCTCGGAGAGAGCATATGACAAAGTCTTCAACACA	988
DB	1476	AGAAGCTGATTGAGCAGTATGAGCTCGGAGAGAGCATATGACAAAGTCTTCAACACA	1417
QY	989	AGGACCTACACAGCAGCTGCTGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGG	1048
DB	1416	AGGACCTACACAGCAGCTGCTGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGG	1357
QY	1049	AGGCAGAAGAGCGGCCACCGAGGAGAGGATTTCTCTGAAAGGCGAGTAGATGCC	1108
DB	1356	AGGCAGAAGAGCGGCCACCGAGGAGAGGATTTCTCTGAAAGGCGAGTAGATGCC	1297
QY	1109	AGAGGATGTGCTGATGATGACACAGACAGACCCACCTGAGCAGACAGCTTGCCTAT	1168
DB	1296	AGAGGATGTGCTGATGATGACACAGACAGACCCACCTGAGCAGACAGCTTGCCTAT	1237
QY	1169	ACACAGAAGAGTTTGAGGAGTTCCAGAACACACTTTTCCAAAAGCAGCGAGGTATTCA	1228
DB	1236	ACACAGAAGAGTTTGAGGAGTTCCAGAACACACTTTTCCAAAAGCAGCGAGGTATTCA	1177
QY	1229	CATTCAAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGCTGGAAGAGAACCA	1288
DB	1176	CATTCAAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGCTGGAAGAGAACCA	1117
QY	1289	CCATGTACCGCTCCCGTGGGAGCAGCAGCAACAGGCCCTGCTTGCATGGCTGAGGAGA	1348
DB	1116	CCATGTACCGCTCCCGTGGGAGCAGCAGCAACAGGCCCTGCTTGCATGGCTGAGGAGA	1057
QY	1349	AAACAGTCCGGGATAAAGAACTGAGGCGCTCAGGTAAAAATCCAAACGGCTGGAGAGC	1408

DB	1056	AAACAGTCCGGGATAAAGAACTGAGGCGCTCAGGTAAAAATCCAAACGGCTGGAGAGC	997
QY	1409	TGTGCGGGCACTGCACAGAGCGCAATGCACTGACAGAGGTTACAGGACCTGAGTG	1468
DB	996	TGTGCGGGCACTGCACAGAGCGCAATGCACTGACAGAGGTTACAGGACCTGAGTG	937
QY	1469	CTGTGTGCCAGGGCTCCCTCACTGACAGTGGCCCTTGAGAGAGGCCAGAGGGCCCTGGGG	1528
DB	936	CTGTGTGCCAGGGCTCCCTCACTGACAGTGGCCCTTGAGAGAGGCCAGAGGGCCCTGGGG	877
QY	1529	CTAAGCACCAGCTCCCGCAGGGTACAGAGGCTTGTCTACCCAGAGGACCCAGCA	1588
DB	876	CTAAGCACCAGCTCCCGCAGGGTACAGAGGCTTGTCTACCCAGAGGACCCAGCA	817
QY	1589	CAGAAGCATCAGCCAGCTGAGGCTTCAAGAGGCCACCTCCCGCAGGGCC	1638
DB	816	CAGAAGCATCAGCCAGCTGAGGCTTCAAGAGGCCACCTCCCGCAGGGCC	767

RESULT 14
 ABZ97034/C
 ID ABZ97034 standard; DNA; 1793 BP.
 XX AC ABZ97034;
 XX DT 17-OCT-2003 (first entry)
 XX Human nucleic acid sequence.
 XX Human; antisense; lung dysfunction; nasal airway dysfunction;
 XX antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
 XX antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
 XX antisense gene therapy; respiratory; lung; adenosine sensitivity;
 XX adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
 XX lung inflammation; respiratory disease; ds.
 XX Homo sapiens.
 OS WO200285308-A2.
 FN 31-OCT-2002.
 PD 23-APR-2002; 2002WO-US013135.
 PF 24-APR-2001; 2001US-0286137P.
 XX (EPIG-) EPIGENESIS PHARM INC.
 XX NYCE JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
 PI Miller S, Tang L, Shahabuddin S;
 XX WPI; 2003-229219/22.
 DR Pharmaceutical composition for treating ailments associated with impaired
 XX respiration, has oligo(s) antisense to specific gene(s) or its
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
 PT ubiquinone.
 XX Disclosure; SEQ ID NO 12276; 872pp; English.
 PS The invention relates to a novel pharmaceutical composition, which has a
 CC first active agent comprising an oligonucleotide antisense to the
 CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
 CC junctions of genes encoding a polypeptide associated with lung and/or
 CC nasal airway dysfunction and a second active agent comprising an
 CC antiinflammatory steroid and ubiquinone. A composition of the invention
 CC has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
 CC immunosuppressive, and cytostatic activity. The composition may have a
 CC use in antisense gene therapy. The composition is useful for treating or
 CC preventing a respiratory, lung or malignant disease or condition, also
 CC for enhancing the prophylactic or therapeutic respiratory effect of an
 CC antiinflammatory steroid in a subject, for reducing or depleting levels

CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;
Query Match 54.3%; Score 890; DB 7; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 749 GGCACAAACCGCTCCCTCAAGAGAGAGAGTGTGCAGCGGCCCGGAGGAGAGAGAGC 808
DB 1656 GGCACAAACCGCTCCCTCAAGAGAGAGAGTGTGCAGCGGCCCGGAGGAGAGAGC 1597
QY 809 GCAAGAGAGTACCTCGCATCTCCAGTTCGACTGAATGACATTCAGCTGCAGATGGAAC 868
DB 1596 GCAAGAGAGTACCTCGCATCTCCAGTTCGACTGAATGACATTCAGCTGCAGATGGAAC 1537
QY 869 AGCACAATGAGCGCAACTCCAGCTCGCCCAAGAGAAATGAGCTGGCTGAGAGGCTCA 928
DB 1536 AGCACAATGAGCGCAACTCCAGCTCGCCCAAGAGAAATGAGCTGGCTGAGAGGCTCA 1477
QY 929 AGAAGCTGATTGAGCAGTATGAGTTCGCGAGAGCATATGCAAAAGTCTTCAAAACA 988
DB 1476 AGAAGCTGATTGAGCAGTATGAGTTCGCGAGAGCATATGCAAAAGTCTTCAAAACA 1417
QY 989 AGGACTCAACAGCAGCTGGTGGATGCTCCAGCTCCAGCGCCCGAGGAGTCTAAAGG 1048
DB 1416 AGGACTCAACAGCAGCTGGTGGATGCTCCAGCTCCAGCGCCCGAGGAGTCTAAAGG 1357
QY 1049 AGGCGAAGAGCGGCACCAAGCGGAGAGAGGATTTCTCTGAAAGAGGAGTGTAGTCCC 1108
DB 1356 AGGCGAAGAGCGGCACCAAGCGGAGAGAGGATTTCTCTGAAAGAGGAGTGTAGTCCC 1297
QY 1109 AGAGATGTGATGATGATGAGCAGCAGAGACCCACTGAGCAACAGCTTGCCCTAT 1168
DB 1296 AGAGATGTGATGATGATGAGCAGCAGAGACCCACTGAGCAACAGCTTGCCCTAT 1237
QY 1169 ACACAGAGAGTTTTCAGAGATTTCCAGACACACTTTCCAAAGAGCGAGGTATTACCA 1228
DB 1236 ACACAGAGAGTTTTCAGAGATTTCCAGACACACTTTCCAAAGAGCGAGGTATTACCA 1177
QY 1229 CATTCAGCAGAGATGGAAGATGACTTAAGAGATCAAGAGCTGGAGAAAGAACCA 1288
DB 1176 CATTCAGCAGAGATGGAAGATGACTTAAGAGATCAAGAGCTGGAGAAAGAACCA 1117
QY 1289 CCATGTACCGTCCCGTGGAGAGCAGCAACAGGCCCTCTGAGATGCTGAGGAGA 1348
DB 1116 CCATGTACCGTCCCGTGGAGAGCAGCAACAGGCCCTCTGAGATGCTGAGGAGA 1057
QY 1349 AAACAGTCGGGATAAAGAACTGGAGGCGCTGCAGGTAAATCCAAACGGCTGGAGAGC 1408
DB 1056 AAACAGTCGGGATAAAGAACTGGAGGCGCTGCAGGTAAATCCAAACGGCTGGAGAGC 997
QY 1409 TGTGCGGCACTCAGACAGAGCCATGACCTGAAACAGAGGATGAGACCTGAGT 1468
DB 996 TGTGCGGCACTCAGACAGAGCCATGACCTGAAACAGAGGATGAGACCTGAGT 937
QY 1469 CTGGTGGCCAGGCGCTCCCTCACTGACGTGGCCCTGAGAGAGGCGCAGAGGCGCTGGG 1528
DB 936 CTGGTGGCCAGGCGCTCCCTCACTGACGTGGCCCTGAGAGAGGCGCAGAGGCGCTGGG 877
QY 1529 CTCAGAGCAGGAGTCCCGGCTCAGAGAGCGCTTGTACCCAGAGCAGCAGCA 1588
DB 876 CTCAGAGCAGGAGTCCCGGCTCAGAGAGCGCTTGTACCCAGAGCAGCAGCA 817
QY 1589 CAGAGAGCATCAGGAGTGGGCTCAGAGAGCGCTTGTACCCAGAGCAGCAGC 1638
DB 816 CAGAGAGCATCAGGAGTGGGCTCAGAGAGCGCTTGTACCCAGAGCAGCAGC 767

RESULT 15

ABZ97033/C

ID ABZ97033 standard; DNA; 1793 BP.

XX AC ABZ97033;

XX AC ABZ97033;

DT 17-OCT-2003 (first entry)

XX Human nucleic acid sequence.

XX Human; antisense; lung dysfunction; nasal airway dysfunction;
XX antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
XX antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
XX antisense gene therapy; respiratory; lung; adenosine sensitivity;
XX adenosine receptor; bronchodilation; lung inflammation; bronchoconstriction; lung allergy;
XX lung inflammation; respiratory disease; ds.

XX Homo sapiens.

XX WO200285308-A2.

XX 31-OCT-2002.

XX 23-APR-2002; 2002WO-US013135.

XX 24-APR-2001; 2001US-0286137P.

XX (EPIG-) EPIGENESIS PHARM INC.

XX Nyce JW, Li Y, Sandrasgra A, Katz E, Pabalan J, Aguilar D;
XX Miller S, Tang L, Shahabuddin S;

XX WPI; 2003-229219/22.

XX Pharmaceutical composition for treating ailments associated with impaired
XX respiration, has oligo(s) antisense to specific gene(s) or its
XX corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
XX ubiquinone.

XX Disclosure; SEQ ID NO 12275; 872pp; English.

XX The invention relates to a novel pharmaceutical composition, which has a
XX first active agent comprising an oligonucleotide antisense to the
XX initiation codon, coding region, 5' or 3' end genomic flanking regions,
XX 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
XX junctions of genes encoding a polypeptide associated with lung and/or
XX nasal airway dysfunction and a second active agent comprising an
XX antiinflammatory steroid and ubiquinone. A composition of the invention
XX has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
XX immunosuppressive, and cytostatic activity. The composition may have a
XX use in antisense gene therapy. The composition is useful for treating or
XX preventing a respiratory, lung or malignant disease or condition, also
XX for enhancing the prophylactic or therapeutic respiratory effect of an
XX antiinflammatory steroid in a subject, for reducing or depleting levels
XX of, or reducing sensitivity to adenosine, reducing levels of adenosine
XX receptor, producing bronchodilation, increasing levels of ubiquinone or
XX lung surfactant in a subject's tissue, or treating bronchoconstriction,
XX lung inflammation, lung allergies, or a respiratory disease or condition.
XX Note: The sequence data for this patent is not represented in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 54.3%; Score 890; DB 7; Length 1793;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 GGCACAAACCGCTCCCTCAAGAGAGAGAGTGTGCAGCGGCCCGGAGGAGAGAGC 808

DB 1656 GGCACAAACCGCTCCCTCAAGAGAGAGAGTGTGCAGCGGCCCGGAGGAGAGAGC 1597

OS	Hom sapiens.	
XX		
XX	Key	Location/Qualifiers
XX	73..1770	
XX	CDS	
XX	/*tag= a	
XX		
XX	W09416074-A2.	
XX		
XX	21-JUL-1994.	
XX		
XX	18-JAN-1994;	94WO-US001101.
XX		
XX	15-JAN-1993;	93US-00005156.
XX		
XX	(UYWA-) UNIV WASHINGTON ST LOUIS.	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	(TEXA) UNIV TEXAS SYSTEM.	
XX		
XX	Ambrus JL, Fauci AS, Ford RJ;	
XX		
XX	WPI; 1994-249221/30.	
XX	P-PSDB; AAR55800.	
XX		
XX	High molecular weight B cell growth factor - able to stimulate B cell	
XX	proliferation and inhibit B cell differentiation, useful to treat e.g.	
XX	systemic lupus erythematosus.	
XX		
XX	Claim 4; Page 75; 95pp; English.	
XX		
XX	The sequence is that of a gene encoding a high molecular weight B cell	
XX	growth factor which stimulates B cell proliferation and inhibits B cell	
XX	differentiation, termed interleukin 14. IL-14 can be used to interfere	
XX	with the replication of B cells and to treat B cell non-Hodgkin's	
XX	lymphoma, systemic lupus erythematosus and systemic necrotizing	
XX	vaculitis. (Updated, on 25-MAR-2003 to correct PN field.)	
XX		
XX	Sequence 1854 BP; 350 A; 559 C; 456 G; 489 T; 0 U; 0 Other;	
XX		

Query March	44.1%	Score 723;	DB 2;	Length 1854;
Best Local Similarity	99.9%;	Prod. No. 0;		
Matches 773;	Conservative	0; Mismatches	1; Indels	0; Gaps
0;				
749 QY	GGCACACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGGCCGGAGGAGGAGGAAGC	808		
1672 Db	GGCACACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGGCCGGAGGAGGAGGAAGC	1613		
809 QY	GCAAGAGGTGACCTCGACATTCAGAGTGACACTGAATGACATTCAGCTGCAGATGGAAC	868		
1612 Db	GCAAGAGGTGACCTCGACATTCAGAGTGACACTGAATGACATTCAGCTGCAGATGGAAC	1553		
869 QY	AGCACAAATGAGCGCAACTCCAAGCTCGCCGAAGAGACATGGAGCTGGCTGAGAGGCTCA	928		
1552 Db	AGCACAAATGAGCGCAACTCCAAGCTCGCCGAAGAGACATGGAGCTGGCTGAGAGGCTCA	1493		
929 QY	AGAGCTGATTTGAGCAGATGATGAGCTCGCGAGAGAGCATATCGACAAAGTCTTCAACACA	988		
1492 Db	AGAGCTGATTTGAGCAGATGATGAGCTCGCGAGAGAGCATATCGACAAAGTCTTCAACACA	1433		
989 QY	AGGACCTTACAACAGCAGCTGTGTGATGCAAGCTCCAGAGCGGCCAGGAGATGCTAAAGG	1048		
1432 Db	AGGACCTTACAACAGCAGCTGTGTGATGCAAGCTCCAGAGCGGCCAGGAGATGCTAAAGG	1373		
1049 QY	AGGCAGAAGACGGCCACAGCGGAGAAAGGATTTTCTCTGAAAGAGCGCAGTATGATGCC	1108		
1372 Db	AGGCAGAAGACGGCCACAGCGGAGAAAGGATTTTCTCTGAAAGAGCGCAGTATGATGCC	1313		
1109 QY	AGGAGTGTGTGCTGATGAGCAGCAGAGAGACCCACCTTGAGCAACAGCTTGCCTAT	1168		
1312 Db	AGGAGTGTGTGCTGATGAGCAGCAGAGAGACCCACCTTGAGCAACAGCTTGCCTAT	1253		
1169 QY	ACACAGAGAAGTTTGAGGAGTTTCCAGAACACACTTCCAAAGACAGCGGATTTACCA	1228		
1252 Db	ACACAGAGAAGTTTGAGGAGTTTCCAGAACACACTTCCAAAGACAGCGGATTTACCA	1193		

QY 1229 CATTCAAGCAGGAGATGAAAAGATGACTTAAGAGATCAAGAGCTGAGAGAAACCA 1288
Db 1192 CATTCAAGCAGGAGATGAAAAGATGACTTAAGAGATCAAGAGCTGAGAGAAACCA 1133
QY 1289 CCATGTACCGTCCCGTGGAGAGCAGCAACAAGCCCTGCTTCCAGATGCTGAGGAGA 1348
Db 1132 CCATGTACCGTCCCGTGGAGAGCAGCAACAAGCCCTGCTTCCAGATGCTGAGGAGA 1073
QY 1349 AAACAGTCCGGGATAAAGAACTGGAGGGCTGCAGGTAAATCAACGGCTGGAGAGC 1408
Db 1072 AAACAGTCCGGGATAAAGAACTGGAGGGCTGCAGGTAAATCAACGGCTGGAGAGC 1013
QY 1409 TGTGCGGGCATCTCAGACAGAGCGCATGCACTGAACAGAGGTACAGACCTGAGTG 1468
Db 1012 TGTGCGGGCATCTCAGACAGAGCGCAATGCACTGAACAGAGGTACAGACCTGAGTG 953
QY 1469 CTGGTGGCCAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGCC 1522
Db 952 CTGGTGGCCAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGCC 899

RESULT 17

AA565026
ID AA565026 standard; cDNA; 5284 BP.

AC AA565026;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #830.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG00839.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 1; SEQ ID NO 830; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (II) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 5284 BP; 1370 A; 1419 C; 1627 G; 868 T; 0 U; 0 Other;

Query Match 41.9%; Score 686; DB 5; Length 5284;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 767 AGAAGAAGGTGTGACGCGGCGCGGAGAGAGAGAAAGCGCAAGAGGTGACCTCGC 826

Db 1310 AGAAGAAGGTGTGACGCGGCGCGGAGAGAGAGAAAGCGCAAGAGGTGACCTCGC 1369

QY 827 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGACAGCAATAGCGGCAACT 886

Db 1370 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGACAGCAATAGCGGCAACT 1429

QY 887 CCAAGCTGCGCCCAAGAGAACATGGAGCTGCTGAGAGGCTCAAGAAGCTGATTGAGCAGT 946

Db 1430 CCAAGCTGCGCCCAAGAGAACATGGAGCTGCTGAGAGGCTCAAGAAGCTGATTGAGCAGT 1489

QY 947 ATGAGCTGCGCGAGGAGCATATCGACAAAGCTTTCAACACAGAGCCTACACAGCAGC 1006

Db 1490 ATGAGCTGCGCGAGGAGCATATCGACAAAGCTTTCAACACAGAGCCTACACAGCAGC 1549

QY 1007 TGGTGGATGCCAAGCTCCAGAGCCCGAGAGATGCTTAAAGAGGCGCAAGAGCGGCAAC 1066

Db 1550 TGGTGGATGCCAAGCTCCAGAGCCCGAGAGATGCTTAAAGAGGCGCAAGAGCGGCAAC 1609

QY 1067 AGCGGAGAGAGGATTTCTCTCTGAAAGAGGAGTAGAGTCCCGAGAGATGTGTGAGCTGA 1126

Db 1610 AGCGGAGAGAGGATTTCTCTCTGAAAGAGGAGTAGAGTCCCGAGAGATGTGTGAGCTGA 1669

QY 1127 TGAAGCAGCAGAGAGAGCCACCTGAGCAACAGCTTGCCCTATACACAGAGAGTTTGAGG 1186

Db 1670 TGAAGCAGCAGAGAGAGCCACCTGAGCAACAGCTTGCCCTATACACAGAGAGTTTGAGG 1729

QY 1187 AGTTCCAGAAACACACTTTTCCAAAAGCAGGAGGTATTCCACCATTTCAACAGCAGAGATGG 1246

Db 1730 AGTTCCAGAAACACACTTTTCCAAAAGCAGGAGGTATTCCACCATTTCAACAGCAGAGATGG 1789

QY 1247 AAAGATGACTAAGAGATCAAGAGCTGAGAAAGAACACCATGTACCGTCCCGGT 1306

Db 1790 AAAGATGACTAAGAGATCAAGAGCTGAGAAAGAACACCATGTACCGTCCCGGT 1849

QY 1307 GGGAGAGCAGCAACAGGCGCTGCTTGGAGATGGCTGAGGAGAAAACAGTCCGGGATAAAG 1366

Db 1850 GGGAGAGCAGCAACAGGCGCTGCTTGGAGATGGCTGAGGAGAAAACAGTCCGGGATAAAG 1909

QY 1367 AACTGGAGGCGCTGACAGTAAAATCCACGGCTGGAGAGAGTGTGCGCGGCACTGCAGA 1426

Db 1910 AACTGGAGGCGCTGACAGTAAAATCCACGGCTGGAGAGAGTGTGCGCGGCACTGCAGA 1969

QY 1427 CAGAGCGCAATGACCTGAACAAGAGG 1452

Db 1970 CAGAGCGCAATGACCTGAACAAGAGG 1995

RESULT 18

AAV88391

ID AAV88391 standard; cDNA; 529 BP.

XX AAV88391;

XX 12-FEB-1999 (first entry)

DE EST clone GP232.

XX 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
XX (BOST-) BOSTON HEART FOUND INC.
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX P-PSDB; AAB82809.
XX New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX Example 4; Fig 24; 143pp; English.
XX The present sequence is that of genomic DNA encoding novel human low
CC density lipoprotein binding protein 3 (LBP-3, see AAB82809). The DNA was
CC isolated from a human genomic library by screening with LBP-3 cDNA (see
CC AAB26501). The open reading frame spans 10 exons. Human LBP-2 nucleic
CC acids are among claimed polynucleotides of the invention that encode
CC novel polypeptides, termed LBPs, capable of binding to native and
CC methylated LDL. Also claimed are isolated LBP polypeptides, and
CC biologically active fragments and analogues of them, as well as
CC expression vectors, cells and methods of producing the LBPs. Methods of
CC determining if an animal is at risk for atherosclerosis, methods for
CC evaluating an agent for use in treating atherosclerosis, and methods for
CC treating a cell having an abnormality in structure or metabolism of LBP
CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
CC nucleic acid, and vaccine compositions, are also claimed
XX Sequence 22255 BP; 5195 A; 5302 C; 5924 G; 5834 T; 0 U; 0 Other;
XX Query Match 20.5%; Score 336; DB 5; Length 22255;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-149;
XX Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCTCAGCCAGACGCTCAGTCTGGGCGCTTCTGATGCTCTGAGAGCTGAGCGCG 231
DB 2999 GCTCAGCCAGACGCTCAGTCTGGGCGCTTCTGATGCTCTGAGAGCTGAGCGCG 3058
QY 232 CAACTGGAAGACATCTGAGCAGATCTGTGGCAATTAACACAGGGGGCCCCGGCGAG 291
DB 3059 CAACTGGAAGACATCTGAGCAGATCTGTGGCAATTAACACAGGGGGCCCCGGCGAG 3118
QY 292 GATGGGCGACGGGTGAGCGGCTGACCGGAGATGACAGAGTCCCGGACCTATGTG 351
DB 3119 GATGGGCGACGGGTGAGCGGCTGACCGGAGATGACAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCTGACCACTCCAGTAGTCAATGGAGAGAAACCTCCCAAG 411
DB 3179 GCAAGGAATGGGAGCTGACCACTCCAGTAGTCAATGGAGAGAAACCTCCCAAG 3238
QY 412 GGGATCCCAACACAGAGAGATCCGGCAGAGTACGAGGTGGAGACCGGACCATCGA 471
DB 3239 GGGATCCCAACACAGAGAGATCCGGCAGAGTACGAGGTGGAGACCGGACCATCGA 3298
QY 472 AGGCCACAG 507
DB 3299 AGGCCACAG 3334
RESULT 20
ACD13448
ID ACD13448 standard; DNA; 115756 BP.
XX ACD13448;
AC AC
XX 13-AUG-2003 (first entry)
DT 13-AUG-2003 (first entry)
XX Human DNA encoding a p53 modifier, SEQ ID 68.
DE Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;
XX Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;
KW

KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
OS Homo sapiens.
XX WO200299122-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017382.
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI WPI; 2003-156859/15.
XX P-PSDB; ABO07223.
XX Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX Example 2; Page 320-351; 678pp; English.
XX The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of a cell (comprising
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence is an HM
XX nucleic acid encoding a p53 pathway modifying protein
XX Sequence 115756 BP; 29575 A; 26665 C; 27977 G; 31539 T; 0 U; 0 Other;
QY Query Match 20.5%; Score 336; DB 7; Length 115756;
DB Best Local Similarity 100.0%; Pred. No. 3.2e-149;
DB Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCTCAGCCAGACGCTCAGTCTGGGCGCTTCTGATGCTCTGAGAGCTGAGCGCG 231
DB 70572 GCTCAGCCAGACGCTCAGTCTGGGCGCTTCTGATGCTCTGAGAGCTGAGCGCG 70631
QY 232 CAACTGGAAGACATCTGAGCAGATCTGTGGCAATTAACACAGGGGGCCCCGGCGAG 291
DB 70632 CAACTGGAAGACATCTGAGCAGATCTGTGGCAATTAACACAGGGGGCCCCGGCGAG 70691

QY 292 GATGGGACAGAGGTGAGCGGCTGACCCGAGATGCCAGAGTCCCGGACCTATGTG 351
Db 70692 GATGGGACAGAGGTGAGCGGCTGACCCGAGATGCCAGAGTCCCGGACCTATGTG 70751
QY 352 GCAAGGAATGGGAGCGCTGACCAACTCCAGTAGTCAATGAGAGAGGAAGCAACCTCCAAAG 411
Db 70752 GCAAGGAATGGGAGCGCTGACCAACTCCAGTAGTCAATGAGAGAGGAAGCAACCTCCAAAG 70811
QY 412 GGGGATCCAAACACAGAGATCCGCGAGTGCAGAGTCCGAGACCGAGACCATCGA 471
Db 70812 GGGGATCCAAACACAGAGATCCGCGAGTGCAGAGTCCGAGACCGAGACCATCGA 70871
QY 472 AGGCCACAGAGAGAGAAAAAGCAAGGGTTTGGGT 507
Db 70872 AGGCCACAGAGAGAGAAAAAGCAAGGGTTTGGGT 70907

RESULT 21
AAS35582
ID AAS35582 standard; cDNA; 405 BP.
XX
AC AAS35582;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human cardiovascular system antigen cDNA polynucleotide SEQ ID NO 467.
XX
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nocotropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasia;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200155321-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001340.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180282P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225216P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.

CC at ftp.wipo.int/pub/published_pct_sequences

Query Match 8.5%; Score 140; DB 4; Length 405;
 Best Local Similarity 100.0%; Pred. No. 5.5e-56;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	230	GCCAACTGGGAAGACATAC	TACTGAGCACATAC	TACTGTTGGACAATAAAC	CAGGGGGCCCCGGCG	289
DB	88	GCCAACTGGGAAGACATAC	TACTGAGCACATAC	TCTGTGGACAATAAAC	CAGGGGGCCCCGGCG	147
QY	290	AGGATGGGGCACAGGGTG	AGCCGGCTGAACCG	AGATGCAGAGAAGT	CCCGACCTATG	349
DB	148	AGGATGGGGCACAGGGTG	AGCCGGCTGAACCG	AGATGCAGAGAAGT	CCCGACCTATG	207
QY	350	TGGCAAGGAATGGGGAGC	CT 369			
DB	208	TGGCAAGGAATGGGGAGC	CT 227			

RESULT 22
 ADE45661
 ID ADE45661 standard; cDNA; 405 BP.
 XX ADE45661;
 XX AC
 XX DT
 XX 29-JAN-2004 (first entry)
 DE Human cardiovascular system related polynucleotide #457.
 XX Human; cardiovascular system related polypeptide; cancer;
 KW proliferative disorder; foetal abnormality; developmental abnormality;
 KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotonic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; gene; ss.
 XX Homo sapiens.
 OS US2003059908-A1.
 XX 27-MAR-2003.
 XX 07-MAR-2002; 2002US-00091504.
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209457P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 28-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-743766/70.

P-PSDB; ADE46276.

New cardiovascular system related polynucleotides and polypeptides, useful for preventing, treating, or ameliorating a medical condition, such as cancer of cardiovascular tissues and cancer metastases.

Claim 1; SEQ ID NO 467; 262pp; English.

The invention relates to human cardiovascular system related polypeptides and the polynucleotides encoding them. The polypeptides, polynucleotides and antibodies to the polypeptides are useful for diagnosing a pathological condition or a susceptibility to a pathological condition, such as for preventing, treating, or ameliorating a medical condition, such as cancer of cardiovascular system tissues, proliferative disorders, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g. rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders and infections. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping. The polypeptides and polynucleotides may also be used as food additives or preservatives to increase or decrease storage capabilities, fat content or other nutritional components. This sequence represents a human cardiovascular system related polynucleotide of the invention.

SQ Sequence 405 BP; 123 A; 90 C; 126 G; 63 T; 0 U; 3 Other;
Query Match 8.5%; Score 140; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.5e-56;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 GCACACTGAGACATCTAGCACATCTGTGTGGACATTAACACAGGGGGCCCGCG 289
DB 88 GCCAACTGGAACATCTAGCACATCTGTGTGGACATTAACACAGGGGGCCCGCG 147

OY 290 AGGATGGGACAGGCTGAGCGGTGACCCGAGAGTCCCGGACCTATG 349
DB 148 AGATGGGACAGGCTGAGCGGTGACCCGAGAGTCCCGGACCTATG 207

OY 350 TGGCAAGGAATGGGAGCCT 369
DB 208 TGGCAAGGAATGGGAGCCT 227

RESULT 23
AAV32836
ID AAV32836 standard; cDNA; 4722 BP.
XX
AC AAV32836;
DT 09-NOV-1998 (first entry)
XX
DE Rabbit low density lipoprotein binding protein LBP-3 cDNA.
XX
KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine; ss.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT CDS 61..1734
FT /*tag= a
FT misc_feature 348..390
FT /*tag= b
FT /*note= "Claim 22"
XX
PN WO9823282-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-US021857.
XX
PR 27-NOV-1996; 96US-0031930P.
PR 03-JUN-1997; 97US-0048547P.
XX
PA (BOST-). BOSTON HEART FOUND INC.
XX
PI Lees AM, Lees RS, Law SW, Arjona AA;
XX
DR WPI; 1998-322455/28.
DR P-PSDB; AAV49039.
XX
PT Nucleic acid encoding low density lipoprotein binding proteins and
PT related vectors - transformed cells, proteins, and modulators of binding,
PT useful for treatment and diagnosis of atherosclerosis and for identifying
PT subjects at risk.
XX
XX Claim 7; Fig 14; 47pp; English.
XX
XX This cDNA clone codes for novel rabbit low density lipoprotein (LDL)
XX binding protein LBP-3 (see AAV49039). It was isolated by functional
XX screening of a cDNA library, produced from balloon-endothelialised
XX healing rabbit abdominal aorta mRNA, for clones encoding LBP-3 able to
XX bind both native and methyl LDL. cDNA clones (see AAV32834-39) and
XX encoded rabbit and human LBP-3 (see AAV49037-42) are claimed. An
XX abnormality in an aspect of LBP metabolism or structure is diagnostic of
XX a risk for atherosclerosis. The invention provides methods for
XX determining if an animal is at risk for atherosclerosis (e.g. for

CC prenatal screening); methods for treating atherosclerosis (including gene
CC therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent
CC formation of atherosclerotic plaque; and methods for treating a cell
CC having an abnormality in LBP structure or metabolism. Pharmaceutical and
CC vaccine compositions are also provided, as well as recombinant vectors
CC and host cells used to produce recombinant LBP
XX
SQ Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;
Query Match 3.6%; Score 59; DB 2; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.8e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1396 CGGCTGAGAGCTGTGCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1454
DB 1456 CGGCTGAGAGCTGTGCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 24
AAH26491
ID AAH26491 standard; cDNA; 4722 BP.
XX
AC AAH26491;
DT 12-NOV-2001 (first entry)
XX
DE Rabbit low density lipoprotein binding protein 3 (LBP-3) cDNA.
XX
KW Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
KW ss.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT CDS 25..1734
FT /*tag= a
XX
PN WO200164874-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006356.
XX
PR 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
XX
PA (BOST-). BOSTON HEART FOUND INC.
XX
PI Lees AM, Lees RS, Law SW, Arjona AA;
XX
DR WPI; 2001-565505/63.
DR P-PSDB; AAB82801.
XX
PT New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Claim 5; Fig 14; 143pp; English.
XX
XX The present sequence is that of cDNA encoding novel rabbit low density
XX lipoprotein binding protein 3 (LBP-3, see AAB82801). The cDNA was
XX isolated following screening of a rabbit cDNA library for clones encoding
XX LBP-3 that bound to both native low density lipoprotein (LDL) and methyl
XX LDL. The invention provides claimed polynucleotides encoding novel
XX polypeptides which are capable of binding to native and methylated LDL,
XX the isolated polypeptides, termed LBP-3, and biologically active fragments
XX and analogues of them, as well as expression vectors, cells and methods
XX of producing the LBP-3. Also claimed are methods of determining if an
XX animal is at risk for atherosclerosis, and methods for evaluating an agent
XX for use in treating atherosclerosis, and methods for treating a cell
XX having an abnormality in structure or metabolism of LBP. Pharmaceutical
XX compositions comprising an LBP polypeptide or nucleic acid, and vaccine
XX compositions, are also claimed

KW Low density lipoprotein binding protein 1; LBP-1; LDL; atherosclerosis;
 KW antiarteriosclerotic; therapy; diagnosis; vaccine; ss.
 XX Mammalia.

OS WO200164874-A2.

FN 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

PA Lees AM, Lees RS, Law SW, Arjona AA;

PI WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,

PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Disclosure; Page 10; 143pp; English.

XX The present sequence is that of a polynucleotide encoding a fragment of

CC novel low density lipoprotein binding proteins (LBPs) of the invention

CC (see AAB2797-820). LBPs are capable of binding to native and methylated

CC low density lipoproteins (LDLs). Isolated polynucleotides encoding novel

CC LBPs and their fragments are claimed, as well as expression vectors,

CC cells and methods of producing the LBPs. Methods of determining if an

CC animal is at risk for atherosclerosis, and methods for evaluating an agent

CC for use in treating atherosclerosis, and methods for treating a cell

CC having an abnormality in structure or metabolism of LBP are also claimed,

CC as are pharmaceutical compositions comprising an LBP polypeptide or

CC nucleic acid, and vaccine compositions

XX Sequence 21 BP; 3 A; 4 C; 8 G; 6 T; 0 U; 0 Other;

SQ Query Match 1.3%; Score 21; DB 5; Length 21;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGTGTCCTCTGAGGCTG 225

Db 1 CGTGTGTCCTCTGAGGCTG 21

RESULT 30

ABT22044

ID ABT22044 standard; DNA; 145 BP.

AC ABT22044;

XX 16-APR-2003 (first entry)

DT Breast cancer marker gene SEQ ID No 417.

DE Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;

XX chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;

XX surrogate marker gene; pharmacodynamic marker gene; transgenic animal;

XX human; ds.

XX Homo sapiens.

OS WO200285298-A2.

PN 31-OCT-2002.

XX 19-APR-2002; 2002WO-US012612.

PA (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;

PI WPI; 2003-093053/08.

DR Novel isolated polypeptide encoded by breast cancer marker gene, useful

XX for diagnosing, staging, monitoring, prognosing and treating diseases

XX associated with breast cancer.

XX Disclosure; Page 147; 725pp; English.

XX The invention relates to an isolated polypeptide encoded by a breast

CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,

CC given in the specification. The methods of the invention are useful for

CC diagnosing patients having an identified breast mass or symptoms

CC associated with breast cancer, to diagnose breast cancer or its

CC precursors, and for monitoring the efficacy of treatment of a breast

CC cancer patient (e.g. efficacy of chemotherapy). The methods are also

CC useful for evaluating a patient before, after or during therapy, to

CC evaluate the reduction in a tumour burden. The breast cancer marker gene

CC proteins are useful as immunogens for raising antibodies, by immunising a

CC mammal with a breast cancer marker protein. The marker proteins are

CC useful as bait proteins in a two-hybrid or three-hybrid assay, to

CC identify other proteins which bind to or interact with the marker

CC proteins. The breast cancer marker genes are useful as surrogate marker

CC genes for one or more disorders, disease states or conditions leading to

CC disease states, in particular, breast cancers. The breast cancer marker

CC genes are useful as pharmacodynamic marker genes. An antibody which

CC selectively binds to a protein of a breast cancer marker gene is useful

CC for treating cancers, particularly breast cancers. The host cell of the

CC invention is useful for producing non-human transgenic animals. This

CC polynucleotide sequence represents one of the breast cancer marker genes

CC of the invention

XX Sequence 145 BP; 29 A; 42 C; 52 G; 19 T; 0 U; 3 Other;

SQ Query Match 1.3%; Score 21; DB 7; Length 145;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGCAGC 147

Db 39 GCAGAGGTCCTCCGCGCAGCAGC 59

RESULT 31

ABV03524/c

ID ABV03524 standard; cDNA; 313 BP.

AC ABV03524;

XX 13-SEP-2002 (first entry)

DT Human prostate expression marker cDNA 3515.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

OS WO200160860-A2.

PN 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 638; 11750pp; English.
PS The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX Sequence 313 BP; 46 A; 103 C; 88 G; 68 T; 0 U; 8 Other;
SQ

Query Match 1.3%; Score 21; DB 5; Length 313;
Best Local Similarity 100.0%; Pred.No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 GCAGAGGTCCCGCAGCAGC 147
DB 182 GCAGAGGTCCCGCAGCAGC 162

RESULT 32
ABV33823/c
ID ABV33823 standard; cDNA; 317 BP.
XX
AC ABV33823;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 33814.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gens; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT

PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 7137; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX Sequence 317 BP; 48 A; 102 C; 93 G; 74 T; 0 U; 0 Other;
SQ

Query Match 1.3%; Score 21; DB 5; Length 317;
Best Local Similarity 100.0%; Pred.No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 GCAGAGGTCCCGCAGCAGC 147
DB 186 GCAGAGGTCCCGCAGCAGC 166

RESULT 33
ABV49859
ID ABV49859 standard; cDNA; 343 BP.
XX
AC ABV49859;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 49850.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 9721; 11750pp; English.
PS The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 343 BP; 74 A; 96 C; 103 G; 70 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 5; Length 343;

Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 37 GCAGAGGTCCTCCGGCAGCAGC 57

RESULT 34

ABK46030/C

ID ABK46030 standard; cDNA; 393 BP.

AC ABK46030;

DT 05-JUN-2002 (first entry)

DE cDNA encoding colon tumour protein, SEQ ID No 1581.

XX Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;
 KW gene; ss.

OS Homo sapiens.

XX WO200212328-A2.

PN 14-FEB-2002.

XX 31-JUL-2001; 2001WO-US024218.

XX 03-AUG-2000; 2000US-0223283P.

PR 28-MAR-2001; 2001US-0279763P.

PR 29-JUN-2001; 2001US-0302051P.

XX (CORI-) CORIXA CORP.

XX King GE, Meagher MJ, Xu J, Secrist H;

XX WPI; 2002-241739/29.

XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
 PT for diagnosing, preventing, and treating colon cancer, and as markers for
 PT the progression of cancer.

XX Claim 1; SEQ ID NO 1581; 147pp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer. ABK4450-
 CC ABK46237 represent coding sequences of human colon tumour proteins of the
 CC invention. Note: With the exception of SEQ ID No 1 and 2, the sequence
 CC data for this patent did not form part of the printed specification but
 CC was supplied by the European Patent Office

XX Sequence 393 BP; 103 A; 95 C; 110 G; 85 T; 0 U; 0 Other;

Query Match

1.3%; Score 21; DB 6; Length 393;

Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 90 GCAGAGGTCCTCCGGCAGCAGC 70

RESULT 35

ABV44401/C

ID ABV44401 standard; cDNA; 444 BP.

XX AC ABV44401;

DT 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 44392.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

PN 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8812; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 444 BP; 82 A; 123 C; 131 G; 108 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 5; Length 444;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 332 GCAGAGGTCCTCCGGCAGCAGC 312

RESULT 36

```
ABV35595/c
ID ABV35595 standard; cDNA; 444 BP.
XX
AC ABV35595;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 35586.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 7402; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 444 BP; 82 A; 123 C; 131 G; 108 T; 0 U; 0 Other;
Query Match 1.3%; Score 21; DB 5; Length 444;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 GCAGAGGTCCTCCGCGAGCAGC 147
DB 332 GCAGAGGTCCTCCGCGAGCAGC 312
RESULT 37
AAS24242/c
ID AAS24242 standard; cDNA; 446 BP.
XX
AC AAS24242;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #423.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003733.
XX
PR 04-FEB-2000; 2000US-0180403P.
PR 28-MAR-2000; 2000US-0192745P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2001-488879/53.
XX
PT New polynucleotides encoding ovarian tumor proteins, useful for treating
PT ovarian cancer, and as probes, primers, and markers of cancer
PT progression.
XX
PS Example 1; Page 180; 378pp; English.
XX
CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated T-
CC cells allowing them to proliferate, and administering to the patient. The
CC sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein cDNA clones
XX
SQ Sequence 446 BP; 107 A; 105 C; 129 G; 105 T; 0 U; 0 Other;
Query Match 1.3%; Score 21; DB 4; Length 446;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 GCAGAGGTCCTCCGCGAGCAGC 147
DB 68 GCAGAGGTCCTCCGCGAGCAGC 48
RESULT 38
AAS23958/c
ID AAS23958 standard; cDNA; 447 BP.
XX
AC AAS23958;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #139.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
```

PF 05-FEB-2001; 2001WO-US003733.
 XX 04-FEB-2000; 2000US-0180403P.
 PR 28-MAR-2000; 2000US-0192745P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Mannion J;
 XX
 DR WPI; 2001-488879/53.
 XX
 PT New polynucleotides encoding ovarian tumor proteins, useful for treating
 PT ovarian cancer, and as probes, primers, and markers of cancer
 PT progression.
 XX
 PS Example 1; Page 132; 378pp; English.
 XX
 XX The invention comprises compositions used for the therapy and diagnosis
 CC of ovarian cancer. The compositions comprise one or more ovarian tumor
 CC proteins, their associated polynucleotides, or immunogenic portions of
 CC the proteins. The ovarian tumor polynucleotides and polypeptides are
 CC useful for stimulating and/or expanding T cells specific for a tumor
 CC protein. They are also useful for inhibiting the development of cancer in
 CC a patient with an ovarian tumor DNA or protein by incubating isolated T-
 CC cells allowing them to proliferate, and administering to the patient. The
 CC sequences can be used as markers for cancer, for example, to monitor
 CC ovarian cancer progression. Probes and primers are useful in nucleic acid
 CC hybridization, in detecting the presence of complementary sequences in a
 CC given sample, for preparing mutant species and for preparing other
 CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
 CC represent human ovarian tumor protein cDNA clones
 XX
 SQ Sequence 447 BP; 106 A; 106 C; 127 G; 106 T; 0 U; 2 Other;
 Query Match 1.3%; Score 21; DB 4; Length 447;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 GCAGAGGTCCTCCGCGACGAC 147
 DB 68 GCAGAGGTCCTCCGCGACGAC 48
 RESULT 39
 ABK53976/c
 ID ABK53976 standard; cDNA; 458 BP.
 AC
 XX
 AC ABK53976;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Human head and neck tumour cDNA, SEQ ID No 171.
 XX
 KW Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200212329-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 01-AUG-2001; 2001WO-US024226.
 XX
 XX 03-AUG-2000; 2000US-0223281P.
 PR 16-NOV-2000; 2000US-0249933P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Fan L;
 XX
 DR WPI; 2002-257467/30.
 XX

PT Novel polynucleotide encoding head and neck tumor polypeptides, useful in
 PT pharmaceutical compositions, e.g. vaccines, for treating head and neck
 PT cancers.
 XX
 PS Claim 1; Page 166; 200pp; English.
 XX
 XX The invention relates to an isolated polynucleotide (I) comprising
 CC sequences selected from 273 sequences fully defined in the specification.
 CC (I), including its encoded polypeptide (II), an antibody binding to (II),
 CC a fusion protein comprising (II) and a T-cell population stimulated by
 CC (I) or (II) are useful for stimulating an immune response in a patient
 CC and treating head and neck cancer in a patient. An oligonucleotide (III)
 CC that hybridizes to (I) is useful for determining the presence of cancer
 CC in a patient, by obtaining a biological sample from the patient,
 CC contacting the sample with (III), detecting in the sample an amount of a
 CC polynucleotide that hybridizes to the oligonucleotide, and comparing the
 CC amount of polynucleotide that hybridizes to the oligonucleotides to a
 CC predetermined cut-off value. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis and
 CC treatment of head and neck cancer. ABK53806-ABK54078 represent human head
 CC and neck cancer cDNA sequences of the invention
 XX
 SQ Sequence 458 BP; 119 A; 109 C; 131 G; 93 T; 0 U; 6 Other;
 Query Match 1.3%; Score 21; DB 6; Length 458;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 GCAGAGGTCCTCCGCGACGAC 147
 DB 99 GCAGAGGTCCTCCGCGACGAC 79
 RESULT 40
 AAH82520/c
 ID AAH82520 standard; cDNA; 491 BP.
 XX
 AC AAH82520;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:144.
 XX
 KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
 KW immunogenic; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151513-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001WO-US001575.
 XX
 PR 14-JAN-2000; 2000US-0176722P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA;
 XX
 DR WPI; 2001-425866/45.
 XX
 PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to
 PT treat and diagnose cancers, particularly ovarian cancer.
 XX
 PS Claim 5; Page 88-89; 338pp; English.
 XX
 XX AAH82377 to AAH83878 represent human ovarian tumour-associated
 CC polynucleotide sequences which encode ovarian tumour proteins. The
 CC ovarian tumour protein and polynucleotide sequences have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. The
 CC ovarian tumour proteins and polynucleotides can be used to inhibit the
 CC development of cancer, particularly ovarian cancer. They can also be used

```
CC to diagnose the onset and progression of cancer
XX Sequence 491 BP; 115 A; 121 C; 140 G; 112 T; 0 U; 3 Other;
SQ

Query Match
  1.3%; Score 21; DB 5; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
Db 68 GCAGAGGTCCTCCGGCAGCAGC 48

RESULT 41
ABK15963/c
ID ABK15963 standard; cDNA; 506 BP.
XX
AC ABK15963;
XX
XX 14-MAR-2002 (first entry)
XX
XX Human lung tumour protein encoding cDNA #75.
XX
XX Human; lung tumour protein; lung cancer; immunostimulant; cytostatic; ss;
XX PCR primer.
XX
XX Homo sapiens.
XX
XX WO200202623-A2.
XX
XX 10-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US020975.
XX
XX 29-JUN-2000; 2000US-0215696P.
XX
XX 22-AUG-2000; 2000US-0227142P.
XX
XX 06-SEP-2000; 2000US-0230481P.
XX
XX 21-DEC-2000; 2000US-0257729P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, McNeill PD, Watanabe Y, Carter D, Henderson RA;
XX Kalos MD;
XX
XX WPI; 2002-106602/14.
XX
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response.
XX
XX Claim 1; Page 155; 277pp; English.
XX
XX The invention relates to polynucleotides encoding lung tumour
XX polypeptides. Compositions comprising the lung tumour polypeptides,
XX polynucleotides, fusion proteins, antibodies to the polypeptides, T cell
XX populations, or antigen presenting cells that express the lung tumour
XX polypeptides, are useful for treating lung cancer or stimulating an
XX immune response. The polynucleotides and polypeptides are also useful in
XX the diagnosis and monitoring of lung cancer. Sequences ABK15889-ABK16343
XX represent cDNA molecules encoding the polypeptides of the invention and
XX PCR primers used to amplify the cDNA
XX
XX Sequence 506 BP; 130 A; 125 C; 141 G; 106 T; 0 U; 4 Other;
XX
XX Query Match 1.3%; Score 21; DB 6; Length 506;
XX Best Local Similarity 100.0%; Pred. No. 22;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
Db 95 GCAGAGGTCCTCCGGCAGCAGC 75

RESULT 42
ABK15963/c
ID ABK15963 standard; cDNA; 506 BP.
XX
AC ABK15963;
XX
XX 25-SEP-2001 (first entry)
XX
XX Human ovarian tumour associated polynucleotide sequence SEQ ID NO:447.
XX
XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
XX immunogenic; vaccine; ss.
XX
XX Homo sapiens.
XX
XX WO200151513-A2.
XX
XX 19-JUL-2001.
XX
XX 16-JAN-2001; 2001WO-US001575.
XX
XX 14-JAN-2000; 2000US-0176722P.
XX
```

PA (CORI-) CORIXA CORP.

PI Algate PA;

XX WPI; 2001-425866/45.

XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to treat and diagnose cancers, particularly ovarian cancer.

PS Claim 5; Page 146-147; 338pp; English.

XX AAH82377 to AAH83878 represent human ovarian tumour-associated polynucleotide sequences which encode ovarian tumour proteins. The ovarian tumour protein and polynucleotide sequences have cytostatic activity, and can be used in gene therapy and vaccine production. The ovarian tumour proteins and polynucleotides can be used to inhibit the development of cancer, particularly ovarian cancer. They can also be used to diagnose the onset and progression of cancer

SQ Sequence 536 BP; 125 A; 134 C; 152 G; 123 T; 0 U; 2 Other;

Query Match 1.3%; Score 21; DB 5; Length 536;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147

Db 68 GCAGAGGTCCTCCGCGACGAC 48

RESULT 44

ACH44414/c

ID ACH44414 standard; cDNA; 538 BP.

XX ACH44414;

XX 13-OCT-2003 (first entry)

XX Human foetal brain cDNA #5139.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.

PS Claim 1; SEQ ID NO 31626; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversities, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030073623

SQ Sequence 538 BP; 135 A; 128 C; 152 G; 113 T; 0 U; 10 Other;

Query Match 1.3%; Score 21; DB 8; Length 538;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147

Db 132 GCAGAGGTCCTCCGCGACGAC 112

RESULT 45

AAC00096/c

ID AAC00096 standard; cDNA; 539 BP.

XX AAC00096;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 94.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00090.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 94; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors

XX Sequence 539 BP; 131 A; 143 C; 152 G; 110 T; 0 U; 3 Other;

Query Match 1.3%; Score 21; DB 3; Length 539;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147

DB 135 GCAGAGGTCCTCCGCGCAGC 115

RESULT 46

AAS56543/C

ID AAS56543 standard; cDNA; 608 BP.

XX

AC AAS56543;

XX 18-DEC-2001 (first entry)

XX Human cDNA for an ovarian cancer protein #167.

XX Human; ss; ovarian cancer protein; cancer; tumour; ovarian cancer;

XX endometrial cancer; cytostatic.

XX Homo sapiens.

XX WO200170976-A2.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US009062.

XX 21-MAR-2000; 2000US-0190710P.

XX 22-JUN-2000; 2000US-0213748P.

XX 19-DEC-2000; 2000US-0257276P.

XX (CORI-) CORIXA CORP.

XX Xu J, Pyle RA, Stolk JA;

XX WPI; 2001-607531/69.

XX Nucleic acids encoding 222 polypeptides associated with ovarian and

XX endometrial cancers, useful for diagnosing, preventing and treating

XX cancers.

XX Claim 1; Page 168; 187pp; English.

XX The invention relates to human polynucleotides encoding proteins
XX associated with ovarian and endometrial cancers. The polynucleotides and
XX the proteins they encode may be used in the prevention, diagnosis and
XX treatment of diseases associated with the inappropriate expression of
XX ovarian and endometrial cancer polypeptides (OECs). For example, the
XX polynucleotide (or an expression vector comprising the polynucleotide)
XX and the OEC may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of OECs by expressing inactive proteins or to
XX supplement the patients own production of them. Additionally, the
XX polynucleotide may be used to produce the OECs, by inserting the nucleic
XX acids into a host cell and culturing the cell to express the protein. The
XX polynucleotide and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acids in samples, and therefore which patients may be in
XX need of restorative therapy. The OECs may also be used as antigens in
XX the production of anti-OEC antibodies and in assays to identify
XX modulators of its expression and activity. The anti-OEC antibodies and
XX antagonists may also be used to down regulate expression and activity.
XX The anti-OEC antibodies may also be used as diagnostic agents for
XX detecting the presence of OEC in samples (e.g. by enzyme linked

CC immunosorbant assay (ELISA)) and hence diagnose patients with cancers.
CC The present sequence is a ovarian and endometrial cancer linked cDNA of
CC the invention

XX Sequence 608 BP; 159 A; 146 C; 178 G; 121 T; 0 U; 4 Other;

Query Match 1.3%; Score 21; DB 4; Length 608;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147

DB 97 GCAGAGGTCCTCCGCGCAGC 77

RESULT 47

ABQ60021/C

ID ABQ60021 standard; cDNA; 657 BP.

XX

AC ABQ60021;

XX 02-AUG-2002 (first entry)

XX Human colon cancer related nucleotide sequence SEQ ID NO:3716.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

XX Genetic analysis; diagnostic; antisense therapy; gene; ss.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US030732.

XX 02-OCT-2000; 2000US-0237271P.

XX (FARB) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;

XX Thiagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer
XX tissues useful for determining the presence of colon cancer in a cell or
XX tissue type, and in antisense therapy.

XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX used in antisense therapy. An antibody immunoreactive with a polypeptide
XX encoded by (I) is useful for detecting cancer in a patient sample, and
XX for detecting the presence or absence of a polynucleotide encoded by a
XX nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX from (I) can be used for determining the presence of a nucleic acid which
XX hybridises to (I), and for determining the phenotype of cells in a sample
XX of cells from a patient. (I) is useful for determining the presence of
XX colon cancer in a cell or tissue type, for determining the presence or
XX state of other type of cancer, in antisense therapy, to generate
XX macroarrays on a solid surface, to identify a chromosome on which the
XX corresponding gene resides, and in tissue profiling, forensics, genetic
XX analysis, mapping and diagnostic applications. (I) can be used to raise
XX antibodies, and to screen for peptide analogues and antagonists

XX Sequence 657 BP; 182 A; 158 C; 152 G; 136 T; 0 U; 29 Other;

Query Match 1.3%; Score 21; DB 6; Length 657;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
 Db |||||
 104 GCAGAGGTCCTCCGCGCAGC 84

RESULT 48

ADE07411/c
 ID ADE07411 standard; DNA; 810 BP.

XX
 AC ADE07411;
 XX

DT 29-JAN-2004 (first entry)
 XX

DE Novel coding sequence (useful for identifying genetic disorders) #477.
 XX

XX novel gene; novel protein; tissue marker; molecular weight marker;
 KW chromosome marker; genetic disorder; gene; ds.
 XX

OS Unidentified.
 XX

PN WO2003054152-A2.
 XX

PD 03-JUL-2003.
 XX

PF 10-DEC-2002; 2002WO-US039555.
 XX

PR 10-DEC-2001; 2001US-0339739P.
 PR

PR 11-DEC-2001; 2001US-0329453P.
 PR

PR 14-MAR-2002; 2002US-0365091P.
 PR

PR 14-MAR-2002; 2002US-0365384P.
 PR

PR 12-APR-2002; 2002US-0372381P.
 PR

PR 12-APR-2002; 2002US-0372615P.
 PR

PR 22-APR-2002; 2002US-00128558.
 PR

PR 24-APR-2002; 2002US-0376045P.
 PR

XX (HYSE-) HYSEQ INC.
 PA

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
 XX

DR WPI; 2003-569235/53.
 DR P-PSDB; ADE08322.

XX New polynucleotides, useful for expressing recombinant proteins for
 PT analysis, characterization or therapeutic use, or as markers for tissues
 PT in which the corresponding protein is preferentially expressed.

XX Claim 1; SEQ ID NO 477; 1177pp; English.
 PS

XX The invention comprises the amino acid and coding sequences of novel
 CC proteins. The DNA and protein sequences of the invention are useful as:
 CC markers for tissues in which the corresponding protein is preferentially
 CC expressed; as molecular weight markers on gels; as chromosome markers or
 CC tags; to identify chromosomes or to map related gene positions; and to
 CC compare with endogenous DNA sequences in patients to identify potential
 CC genetic disorders. The present DNA sequence represents a gene of the
 CC invention.

XX Sequence 810 BP; 185 A; 233 C; 246 G; 146 T; 0 U; 0 Other;
 SQ

Query Match 1.3%; Score 21; DB 9; Length 810;
 Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
 Db |||||
 270 GCAGAGGTCCTCCGCGCAGC 250

RESULT 49

ABK34931/c

ID ABK34931 standard; cDNA; 824 BP.

XX
 AC ABK34931;
 XX

DT 08-MAY-2002 (first entry)
 XX

DE Human cDNA encoding secreted protein #69.
 XX

XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.

XX Homo sapiens.
 OS

PN WO200177288-A2.
 XX

PD 18-OCT-2001.
 XX

PF 29-MAR-2001; 2001WO-US010224.
 XX

PR 06-APR-2000; 2000US-0195582P.
 PR

XX (GEMY) GENETICS INST INC.
 PA

PI Wong GG, Clark HF, Rechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 XX

DR WPI; 2002-179321/23.
 XX

PT Five hundred and ninety two polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for treating
 PT immune deficiencies and disorders such as autoimmune disorders.

XX Claim 1; Page 102-103; 372pp; English.
 PS

XX The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention

XX Sequence 824 BP; 204 A; 215 C; 240 G; 165 T; 0 U; 0 Other;
 SQ

Query Match 1.3%; Score 21; DB 6; Length 824;
 Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
 Db |||||
 97 GCAGAGGTCCTCCGCGCAGC 77

RESULT 50

AAA34999/c

ID AAA34999 standard; DNA; 833 BP.

XX

AAA34999;
 28-JUL-2000 (first entry)
 Human adenosine receptor related polynucleotide SEQ ID NO:2688.

Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

Homo sapiens.
 WO200009525-A2.
 24-FEB-2000.
 03-AUG-1999; 99WO-US017712.
 03-AUG-1998; 98US-0095212P.
 (UYEC-) UNIV EAST CAROLINA.
 Nyce JW;
 WPI; 2000-205971/18.

New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers.

Disclosure; Page 898; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation. Impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1880 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing

Sequence 833 BP; 215 A; 204 C; 238 G; 176 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 3; Length 833;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

127 GCAGAGGTCGCGGACGC 147
 69 GCAGAGGTCGCGGACGC 49

RESULT 51
 AAF21121/c
 ID AAF21121 standard; DNA; 833 BP.
 XX AC AAF21121;
 XX 14-MAR-2001 (first entry)
 XX Human low adenosine antisense oligonucleotide related sequence #2688.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.
 WO2000062736-A2.
 26-OCT-2000.
 24-MAR-2000; 2000WO-US008020.
 06-APR-1999; 99US-0127958P.
 (UYEC-) UNIV EAST CAROLINA.
 (NYCE/) NYCE J W.
 Nyce JW;
 WPI; 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions.

Disclosure; Page 973; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system peptide receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide

CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention

XX Sequence 833 BP; 215 A; 204 C; 238 G; 176 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 3; Length 833;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147
|||||
Db 69 GCAGAGGTCCTCCGCGACGAC 49

RESULT 52
ABZ96815/c
ID ABZ96815 standard; DNA; 833 BP.

XX AC ABZ96815;

XX DT 17-OCT-2003 (first entry)

XX DE Human nucleic acid sequence.

XX KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiqunone; antiinflammatory; anti-allergic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.

XX OS Homo sapiens.

XX PN WO200285308-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013135.

XX PR 24-APR-2001; 2001US-0286137P.

XX PA (EPIG-) EPIGENESIS PHARM INC.

XX PI Nvce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;

XX DR WPI; 2003-229219/22.

XX PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiqunone.

XX PS Disclosure; SEQ ID NO 12057; 872pp; English.

XX CC The invention relates to a novel pharmacological composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiqunone. A composition of the invention
CC has antiinflammatory, anti-allergic, antiasthmatic, hypotensive,
CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiqunone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 833 BP; 215 A; 204 C; 238 G; 176 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 7; Length 833;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147
|||||
Db 69 GCAGAGGTCCTCCGCGACGAC 49

RESULT 53
ABV22299/c
ID ABV22299 standard; cDNA; 838 BP.

XX AC ABV22299;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 22290.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PS Claim 1; Page 3854; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 838 BP; 194 A; 224 C; 251 G; 169 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 5; Length 838;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
 Db 140 GCAGAGGTCCCGCAGCAGC 120

RESULT 54
 AAA34998/c
 ID AAA34998 standard; DNA; 851 BP.
 AC AAA34998;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human adenosine receptor related polynucleotide SEQ ID NO:2687.
 XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009525-A2.
 XX
 PD 24-FEB-2000.
 XX
 XX 03-AUG-1999; 99WO-US017712.
 PF
 XX 03-AUG-1998; 98US-0095212P.
 PR
 XX (UYEC-) UNIV EAST CAROLINA.
 PA
 XX Nyce JW;
 PI
 XX WPI; 2000-205971/18.
 DR
 XX New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers.
 PS Disclosure; Page 898; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cyrostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1880 (AAA3232 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing

QY Sequence 851 BP; 222 A; 210 C; 240 G; 179 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 3; Length 851;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
 Db 83 GCAGAGGTCCCGCAGCAGC 63

RESULT 55
 AAF21120/c
 ID AAF21120 standard; DNA; 851 BP.
 XX
 AC AAF21120;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human low adenosine antisense oligonucleotide related sequence #2687.
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; ROS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2000062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 XX 24-MAR-2000; 2000WO-US008020.
 PF
 XX 06-APR-1999; 99US-0127958P.
 PR
 XX (UYEC-) UNIV EAST CAROLINA.
 PA
 XX (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 XX WPI; 2000-679539/66.
 DR
 XX Low adenosine (A) content antisense oligonucleotides which do not trigger
 PT adenosine receptors during metabolism, useful e.g. for treating cancers
 PT and respiratory obstructions.
 PS Disclosure; Page 973; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders

CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention
XX
SQ Sequence 851 BP; 222 A; 210 C; 240 G; 179 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 3; Length 851;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGCAGC 147
|||
DB 83 GCAGAGGTCCTCCGCGCAGCAGC 63

RESULT 56
ABZ96814/c
ID ABZ96814 standard; DNA; 851 BP.
XX
AC ABZ96814;
XX
DT 17-OCT-2003 (first entry)
XX
DE Human nucleic acid sequence.
XX
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiqunone; antiinflammatory; anti-allergic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
OS Homo sapiens.
XX
PN WO200285308-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013135.
XX
PR 24-APR-2001; 2001US-0286137P.
XX
PA (EPIC-) EPTGENESIS PHARM INC.
XX

NYce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
DR WPI; 2003-229219/22.
XX
PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiqunone.
XX
PS Disclosure; SEQ ID NO 12056; 872pp; English.
XX

CC The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiqunone. A composition of the invention
CC has antiinflammatory, anti-allergic, antiasthmatic, hypotensive, and
CC immunosuppressive, and cytostatic activity. The composition may have a

CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiqunone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 851 BP; 222 A; 210 C; 240 G; 179 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 7; Length 851;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGCAGC 147
|||
DB 83 GCAGAGGTCCTCCGCGCAGCAGC 63

RESULT 57
ADB31348/c
ID ADB31348 standard; cDNA; 851 BP.
XX
AC ADB31348;
XX
DT 04-DEC-2003 (first entry)
XX
DE Bicalutamide regulated prostate cancer gene #25.
XX
KW Human; ss; prostate cancer; cancer; androgen; castration;
KW anti-androgenic drug; bicalutamide; casodex; testosterone.
XX
OS Homo sapiens.
XX
PN US6506607-B1.
XX
PD 14-JAN-2003.
XX
PF 23-DEC-1998; 98US-00220132.
XX
PR 24-DEC-1997; 97US-0068821P.
PR 25-MAR-1998; 98US-0079303P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Shyjan AW;
XX
DR WPI; 2003-595981/56.
XX

Determination of whether compound is useful for prostate cancer treatment
PT comprises measuring expression level of specific nucleic acid sequence in
PT prostate cancer cell sample in the presence and absence of compound.
XX

Example 2; Col 197-200; 194pp; English.

CC The invention discloses a method for determining whether a compound can
CC be used to treat prostate cancer and comprises measuring the expression
CC level of a nucleic acid in prostate cancer cell sample in the presence
CC and absence of the compound. Determining whether a compound can be used
CC to treat prostate cancer comprises identifying the compound as useful for
CC prostate cancer treatment when the expression level of the nucleic acid
CC in the presence of the compound is less than the expression level of the
CC nucleic acid in the absence of the compound. Prostate cancer is usually
CC treated by androgen withdrawal, by castration or through the use of an
CC anti-androgenic drug. Bicalutamide (casodex) is one such anti-androgenic
CC compound. The determination whether a selected compound, e.g. an anti-
CC androgenic compound or testosterone, can be used to treat the prostate
CC cancer can be made on a patient by patient basis. The sequence presented
CC is a gene which is more highly expressed in bicalutamide treated prostate

CC cancer cells than the untreated cells.

SQ Sequence 851 BP; 222 A; 210 C; 240 G; 179 T; 0 U; 0 Other;
Query Match 1.3%; Score 21; DB 9; Length 851;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGGAGCAGC 147
DB 83 GCAGAGGTCCCGGAGCAGC 63

RESULT 58
ABK34930/C
ID ABK34930 standard; cDNA; 857 BP.
XX
AC ABK34930;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human cDNA encoding secreted protein #68.
XX
KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
KW tissue regeneration; wound healing; burn; haematopoiesis;
KW myeloid cell deficiency; lymphoid cell deficiency.
XX
OS Homo sapiens.
XX
PN WO200177288-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010224.
XX
PR 06-APR-2000; 2000US-0195582P.
XX
PA (GEM) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
XX
DR WPI; 2002-179321/23.
XX
PT Five hundred and ninety two polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for treating
PT immune deficiencies and disorders such as autoimmune disorders.
XX
PS Claim 1; Page 102; 372pp; English.
XX
CC The invention relates to 592 polynucleotides which have been derived from
CC a variety of human tissue sources and which encode novel secreted
CC proteins. The polynucleotides can be used as probes for the
CC identification and isolation of full length cDNA and genomic DNA. The
CC polynucleotides and proteins can also be used as nutritional supplements.
CC The proteins are useful in the treatment of various immune deficiencies
CC and disorders such as viral infections, bacterial infections, fungal
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
CC and conditions (e.g. asthma). They are also useful for treating
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
CC useful for tissue regeneration, for wound healing and in the treatment of
CC burns, incisions and ulcers. The proteins are also useful for regulating
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention

SQ Sequence 857 BP; 214 A; 215 C; 246 G; 182 T; 0 U; 0 Other;
Query Match 1.3%; Score 21; DB 6; Length 857;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGGAGCAGC 147
DB 97 GCAGAGGTCCCGGAGCAGC 77

RESULT 59
ABK84057/C
ID ABK84057 standard; cDNA; 893 BP.
XX
AC ABK84057;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #628.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 628; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA) by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an

CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 893 BP; 225 A; 228 C; 250 G; 190 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 6; Length 893;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 GCAGAGGTCCTCCGCGACGAC 147
 Db 116 GCAGAGGTCCTCCGCGACGAC 96
 RESULT 60
 ABK64544/c
 ID ABK64544 standard; DNA; 893 BP.
 XX
 XX
 AC ABK64544;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human benign prostatic hyperplasia gene #439.
 XX
 KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO200212440-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 07-AUG-2001; 2001WO-US024708.
 XX
 PR 07-AUG-2000; 2000US-0223323P.
 PR 05-JUN-2001; 2001US-00873319.
 XX
 XX (GENE-) GENE LOGIC INC.
 PA (NISR) JAPAN TOBACCO INC.
 XX
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 XX
 DR WPI; 2002-257476/30.
 XX
 XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells.
 XX
 XX Disclosure; Page 257; 444pp; English.
 XX
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the

CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles. (I)
 CC is useful for diagnosing the onset or progression of BPH. (II) is useful
 CC for identifying an agent that modulates the onset or progression of BPH.
 CC The methods are useful to present information identifying the expression
 CC level in a tissue or cells, by comparing the expression level of genes
 CC given in the specification in the tissue or cells to the level of
 CC expression of gene in the database, and displaying the expression levels
 CC of at least one gene in the tissue or cell sample compared to the
 CC expression level in BPH. Agents using (II) are useful for treating BPH or
 CC prostate cancer. ABK64106-ABK64960 represent human benign prostatic
 CC hyperplasia gene sequences of the invention
 XX
 XX Sequence 893 BP; 225 A; 228 C; 250 G; 190 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 6; Length 893;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 GCAGAGGTCCTCCGCGACGAC 147
 Db 116 GCAGAGGTCCTCCGCGACGAC 96
 RESULT 61
 ABN95879/c
 ID ABN95879 standard; DNA; 893 BP.
 XX
 XX
 AC ABN95879;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #2377 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 FN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US030589.
 XX
 PR 02-OCT-2000; 2000US-0237054P.
 XX
 XX (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX
 DR WPI; 2002-426119/45.
 XX
 XX Diagnosing and detecting the progression of liver cancer, hepatocellular
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the
 PT level of expression of two or more genes in a liver tissue sample.
 XX
 PS Claim 1; SEQ ID NO 2377; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer; hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 893 BP; 225 A; 228 C; 250 G; 190 T; 0 U; 0 Other;

XX Query Match 1.3%; Score 21; DB 6; Length 893;
XX Best Local Similarity 100.0%; Pred. NO. 22;

XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

DB 116 GCAGAGGTCCCGCAGCAGC 96

RESULT 62

ABX10431/c

ID ABX10431 standard; DNA; 893 BP.

AC ABX10431;

XX 28-JAN-2003 (first entry)

XX DNA encoding protein differentially regulated in prostate cancer #100.

XX Prostate cancer; gene expression; differential regulation;

XX molecular marker; drug target; cancer detection; cancer diagnosis;

XX cancer staging; cancer grading; cancer assessing; cancer monitoring;

XX gene; ds.

XX Homo sapiens.

XX WO200281638-A2.

XX 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.

XX 06-APR-2001; 2001US-0281731P.

XX 06-APR-2001; 2001US-0281732P.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Jay G;

XX WPI; 2003-058520/05.

XX Novel genes which are differentially regulated in prostate cancer, useful

XX for diagnosing prostate cancer in prostate tissue sample and assessing

XX therapeutic or preventive intervention in prostate cancer patients.

XX Claim 1; Page 192-193; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in

XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a

XX sample comprising prostate tissue, which involves determining the number

XX of target genes which are differentially regulated in the sample, where

XX the number is indicative of the probability that the sample comprises

XX prostate cancer. (I) is useful for assessing a therapeutic or preventive

XX intervention in a subject having a prostate cancer, which involves

XX determining the expression levels in a sample comprising prostate tissue

XX of target genes which are differentially regulated in prostate cancer.

XX Preferably, the expression levels of at least 10 genes are determined.

XX (I) is also useful for identifying agents that modulate a biological

XX activity of a polypeptide differentially regulated in prostate cancer

XX cells, which involves contacting a polypeptide differentially-regulated

XX in prostate cancer cells with a test agent under conditions effective for

XX the test agent to modulate a biological activity of the polypeptide, and

XX determining whether the test agent modulates the biological activity. (I)

XX is useful as molecular markers, as drug targets, and for detecting,

XX diagnosing, staging, grading, assessing, monitoring, prognosticating,

XX preventing or treating, determining predisposition to diseases and

XX conditions especially relating to prostate cancer. (I) and its expression

XX products are used in the diagnostic test to assay for presence of cancer

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is useful
CC in therapeutic applications to treat prostate cancer. The identification
CC of specific genes, and groups of genes, expressed in pathways
CC physiologically relevant to prostate cancer permits the definition of
CC functional and disease pathways and the delineation of targets in these
CC pathways which are useful in diagnostic, therapeutic, and clinical
CC applications. This sequence encodes a protein differentially regulated in
CC prostate cancer

XX Sequence 893 BP; 225 A; 228 C; 250 G; 190 T; 0 U; 0 Other;

XX Query Match 1.3%; Score 21; DB 7; Length 893;

XX Best Local Similarity 100.0%; Pred. NO. 22;

XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

DB 116 GCAGAGGTCCCGCAGCAGC 96

RESULT 63

ACD13434/c

ID ACD13434 standard; cDNA; 893 BP.

XX AC ACD13434;

XX 13-AUG-2003 (first entry)

XX Human DNA encoding a p53 modifier, SEQ ID 106.

XX Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;

XX antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;

XX kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;

XX apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX WO200299122-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-US017382.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-156859/15.

XX P-PSDB; ABC07261.

XX Identifying modulators of the p53 pathway for use in treating apoptotic

XX or cell proliferation disorders, comprises screening for agents that

XX modulate activity of a human ortholog of genes that modify the p53

XX pathway in Drosophila.

XX Example 2; Page 394; 678pp; English.

XX The invention relates to identifying (M1) a candidate p53 pathway

XX modulating agent, by contacting an assay system comprising a purified HM

XX polypeptide (human orthologue of genes that modify the p53 pathway in

XX Drosophila) or nucleic acid with a test agent under conditions, where but

XX for the presence of the test agent, the system provides a reference

XX activity, and detecting a test agent-biased activity of the assay system.

XX Also included are modulating (M2) a p53 pathway of a cell (comprising

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence is an HM
 CC nucleic acid encoding a p53 pathway modifying protein

XX Sequence 893 BP; 225 A; 228 C; 250 G; 190 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 7; Length 893;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGGCAGCAGC 147
 Db 116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 64

AACT78087/c
 ID AAC78087 standard; cDNA; 1120 BP.

AC AACT78087;

XX 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:481.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW anti-diabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW anti-inflammatory; antithyroid; anti-allergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.

XX Homo sapiens.

OS WO20005350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX P-PSDB; AAB43878.

XX Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer.

XX Claim 1; Page 1014-1014; 2352pp; English.

XX AACT77607 to AACT78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC anti-inflammatory; antithyroid; anti-allergic; thrombolytic; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neutropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological diseases and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AACT78449 to
 CC AACT78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention

XX Sequence 1120 BP; 281 A; 317 C; 313 G; 207 T; 0 U; 2 Other;

Query Match 1.3%; Score 21; DB 3; Length 1120;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGGCAGCAGC 147

Db 320 GCAGAGGTCCTCCGGCAGCAGC 300

RESULT 65

ABQ54919/c

ID ABQ54919 standard; cDNA; 1131 BP.

XX ABQ54919;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HTFML39 cDNA, SEQ ID NO:799.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW anti-inflammatory; gynaecological; reproductive; gene; ss.

XX Homo sapiens.

OS WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

DR P-PSDB; ABP41842.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
XX Claim 1; SEQ ID NO 799; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1131 BP; 267 A; 328 C; 321 G; 212 T; 0 U; 3 Other;

Query Match 1.3%; Score 21; DB 6; Length 1131;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 GCAGAGGTCCTCCGGCAGCAGC 147
Db 320 GCAGAGGTCCTCCGGCAGCAGC 300

RESULT 66
AAH72841/c
ID AAH72841 standard; cDNA; 1185 BP.

XX AC AAH72841;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 4115.

XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US033312.

XX PR 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
XX 21-JUL-2000; 2000US-0220114P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX Claim 1; Page 884; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy

SQ Sequence 1185 BP; 309 A; 303 C; 314 G; 248 T; 0 U; 11 Other;

Query Match 1.3%; Score 21; DB 4; Length 1185;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 GCAGAGGTCCTCCGGCAGCAGC 147

Db 292 GCAGAGGTCCTCCGGCAGCAGC 272

RESULT 67

ABV21116/c

ID ABV21116 standard; cDNA; 1307 BP.

XX AC ABV21116;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 21107.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 3493-3494; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1307 BP; 416 A; 330 C; 353 G; 208 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 5; Length 1307;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147

DB 211 GCAGAGGTCCTCCGGCAGCAGC 191

RESULT 68

ABV26958/c

ID ABV26958 standard; cDNA; 1307 BP.

XX AC ABV26958;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 26949.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5461; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX specification or its complement. (I) is useful for: (a) assessing whether

XX a patient is afflicted with prostate cancer; (b) monitoring the

XX progression of prostate cancer in a patient; (c) assessing the efficacy

XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1307 BP; 416 A; 330 C; 353 G; 208 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 5; Length 1307;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147

DB 211 GCAGAGGTCCTCCGGCAGCAGC 191

RESULT 69

AAA35002/c

ID AAA35002 standard; DNA; 10032 BP.

XX AC AAA35002;

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide SEQ ID NO:2691.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;

XX phosphorothioate; impaired respiration; inflammation; allergy;

XX allergic disease; bronchoconstriction; inhibitor; anti-inflammatory;

XX anti-allergic; antiasthmatic; cytostatic; analgesic; impaired airway;

XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;

XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;

XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;

XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US017712.

XX 03-AUG-1998; 98US-0095212P.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary

XX vasoconstriction, inflammation, allergies, asthma, hypertension,

XX bronchitis, emphysema, respiratory distress syndrome, ischemia or

XX cancers.

XX Disclosure; Page 900-903; 1343pp; English.

XX The present invention describes a new composition comprising an antisense

XX oligonucleotide (ON) with low adenosine (up to 15%), which targets

XX nucleic acids involved in bronchoconstriction, allergies, and/or

XX inflammation. The ON can have anti-inflammatory, anti-allergic,

XX antiasthmatic, cytostatic and analgesic activities. The compositions are

XX useful for the treatment of diseases associated with inflammation,

XX impaired airways, including lung disease and diseases whose secondary

XX effects afflict the lungs of a subject. They can be used for treating

XX e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,

XX impeded respiration, respiratory distress syndrome, pain, cystic

XX fibrosis, pulmonary hypertension, emphysema, chronic obstructive

XX pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,

XX carcinomas, and cancers which may metastasize to the lungs, including

XX breast and prostate cancer. The reduction of the adenosine content of the

CC ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA3213 to AAA3512 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 185, and then the last 185
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
 CC AAA33992) are specifically claimed ONs from the present invention. N.B.
 CC Sequences given in the disclosure of the present invention do not match
 CC up with their corresponding SEQ ID NO: sequences given in the sequence
 CC listing
 XX
 SQ Sequence 10032 BP; 2342 A; 2486 C; 2643 G; 2561 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 3; Length 10032;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
 |||||
 DB 1834 GCAGAGGTCCCGCAGCAGC 1814

RESULT 70
 AAF21124/c
 ID AAF21124 standard; DNA; 10032 BP.

XX
 ID AAF21124;
 AC AAF21124;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human low adenosine antisense oligonucleotide related sequence #2691.
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilation; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.

XX Homo sapiens.
 OS
 PN WO200062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US008020.
 XX
 PR 06-APR-1999; 99US-0127958P.
 XX
 PA (UYSC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 XX Low adenosine (A) content antisense oligonucleotides which do not trigger
 PT adenosine receptors during metabolism, useful e.g. for treating cancers
 PT and respiratory obstructions.
 PT
 PS Disclosure; Page 975-978; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'universal' or alternative base.
 CC (I) can have respiratory, bronchodilation, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
 CC surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention
 XX

SQ Sequence 10032 BP; 2342 A; 2486 C; 2643 G; 2561 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 3; Length 10032;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
 |||||
 DB 1834 GCAGAGGTCCCGCAGCAGC 1814

RESULT 71
 ABZ96818/c
 ID ABZ96818 standard; DNA; 10032 BP.

XX
 AC ABZ96818;
 XX
 DT 17-OCT-2003 (first entry)
 XX
 DE Human nucleic acid sequence.
 XX
 KW Human; antisense; lung dysfunction; nasal airway dysfunction;
 KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
 KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
 KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
 KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
 KW lung inflammation; respiratory disease; ds.

OS Homo sapiens.
 XX
 PN WO200285308-A2.
 XX

PD 31-OCT-2002.

PF 23-APR-2002; 2002WO-US013135.

PR 24-APR-2001; 2001US-0286137P.

PA (EPIG-) EPIGENESIS PHARM INC.

XX
 PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
 PI Miller S, Tang L, Shahabuddin S;
 XX

DR WPI; 2003-229219/22.

XX Pharmaceutical composition for treating ailments associated with impaired
 PT respiration, has oligo(s) antisense to specific gene(s) or its
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
 PT ubiquinone.

XX PS Disclosure; SEQ ID NO 12060; 872pp; English.

XX CC The invention relates to a novel pharmaceutical composition, which has a first active agent comprising an oligonucleotide antisense to the initiation codon, coding region, 5' or 3' end genomic flanking regions, 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of junctions of genes encoding a polypeptide associated with lung and/or nasal airway dysfunction and a second active agent comprising an anti-inflammatory steroid and ubiquinone. A composition of the invention has anti-inflammatory, anti-allergic, antiasthmatic, hypotensive, immunosuppressive, and cytostatic activity. The composition may have a use in antisense gene therapy. The composition is useful for treating or preventing a respiratory, lung or malignant disease or condition, also for enhancing the prophylactic or therapeutic respiratory effect of an anti-inflammatory steroid in a subject, for reducing or depleting levels of, or reducing sensitivity to adenosine, reducing levels of adenosine receptor, producing bronchodilation, increasing levels of ubiquinone or lung surfactant in a subject's tissue, or treating bronchoconstriction, lung inflammation, lung allergies, or a respiratory disease or condition. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 10032 BP; 2342 A; 2486 C; 2643 G; 2561 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 7; Length 10032;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 GCAGAGGTCCTCCGCGACGAC 147
|||||
Db 1834 GCAGAGGTCCTCCGCGACGAC 1814

RESULT 72
ABX45116/c
ID ABX45116 standard; cDNA; 328 BP.
XX AC ABX45116;
XX DT 21-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #10281.
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX KW muscle deposition; fat deposition; genome mapping; gene identification;
XX KW gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-00960352.
XX PR 12-JAN-1999; 99US-0115707P.
XX PR 11-JAN-2000; 2000US-00480902.
XX PA (BYATT/) BYATT J C.
XX PA (MATH/) MATHIALAGAN N.
XX PA (TAON/) TAO N.
XX PA (WARRE/) WARREN W C.
XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX DR WPI; 2003-110599/10.
XX PT New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.

PS Claim 2; SEQ ID NO 10281; 245pp; English.

XX CC The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridize to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridization between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site:
seqdata.uspto.gov/sequence.html?DocID=20020137139

XX SQ Sequence 328 BP; 20 A; 127 C; 123 G; 58 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 786 GCCCGGGGAGGAGGAGGAGA 805
|||||
Db 239 GCCCGGGGAGGAGGAGGAGA 220

RESULT 73
ABZ14050/c
ID ABZ14050 standard; DNA; 576 BP.
XX AC ABZ14050;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1855.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US026685.
XX PR 24-AUG-2000; 2000US-0227866P.
XX PR 26-JAN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX PA (Scri) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Kreps J, Wang X, Zhu T;
XX DR WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 1855; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

CC cell has been exposed, comprising: (a) contacting nucleic acid

CC representative of expressed polynucleotides in the plant cell with an

CC array or probes representative of the plant cell genome; and (b)

CC detecting a profile of expressed polynucleotides in the plant cell

CC characteristic of a stress response. The method is useful in the

CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that

CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

CC in methods of the invention. Note: The sequence data for this patent is

CC not represented in the printed specification but is based on sequence

CC information supplied to Derwent by the European Patent Office

XX

SQ Sequence 576 BP; 107 A; 241 C; 80 G; 148 T; 0 U; 0 Other;

Query Match 1-2%; Score 20; DB 6; Length 576;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CGGAGGAGGAGGAGGAGGAGCG 809

Db 289 CGGAGGAGGAGGAGGAGGAGCG 270

RESULT 74

AAC42439/c

ID AAC42439 standard; DNA; 602 BP.

XX AAC42439;

AC AAC42439;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35562.

DE Hybridisation assay; genetic mapping; gene expression control;

XX protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0128264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 14-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0136299P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140951P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144684P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148884P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 23-OCT-1999; 99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 1.2%; Score 20; DB 3; Length 602;

Best Local Similarity 100.0%; Pred. NO. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CGGGAGGAGGAGGAGGAGCG 809

Db 315 CGGGAGGAGGAGGAGGAGCG 296

RESULT 75

AAH07907

ID AAH07907 standard; cDNA; 710 BP.

XX AC AAH07907;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:4742.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 1; SEQ ID NO 4742; 2537bp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

SQ Sequence 710 BP; 224 A; 144 C; 203 G; 136 T; 0 U; 3 Other;

Query Match 1.2%; Score 20; DB 4; Length 710;
 Best Local Similarity 100.0%; Pred. No. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCCAGAGGAGAGCTGGC 569

Db 502 ACCCCAGAGGAGAGCTGGC 521

RESULT 76

ID AAF22469
 AC AAF22469 standard; cDNA; 823 BP.

AC AAF22469;

DT 26-MAR-2001 (first entry)

DE Human breast cancer associated antigen nucleotide sequence SEQ ID NO:48.

KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytostatic; cancer vaccine; ss.

OS Homo sapiens.

PN WO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US014749.

PR 28-MAY-1999; 99US-0136526P.

PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Obata Y;

DR WPI; 2001-025274/03.

PT Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.

PS Claim 50; Page 289-290; 799pp; English.

CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC reduction of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer

SQ Sequence 823 BP; 324 A; 155 C; 192 G; 149 T; 0 U; 3 Other;

Query Match 1.2%; Score 20; DB 4; Length 823;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 550 ACCCCAGAGGAGAGCTGGC 569
 Db 17 ACCCCAGAGGAGAGCTGGC 36

RESULT 77

AAV29064
 ID AAV29064 standard; cDNA; 1191 BP.

AC AAV29064;

DT 28-AUG-1998 (first entry)

DE BRCA1 modulator protein 091-132Q20 cDNA.

KW BRCA1 modulator protein; 091-132Q20; breast cancer antigen 1;
 KW tumour suppressor protein; diagnosis; therapy; human; ss.

OS Homo sapiens.

PH Key Location/Qualifiers
 FT CDS 34..1191
 FT /*tag= a

PN WO9810066-A1.

PD 12-MAR-1998.

PF 06-AUG-1997; 97WO-US013944.

PR 04-SEP-1996; 96US-0025601P.

XX (ONYX-) ONYX PHARM INC.

PI Rubinfeld B, Polakis P, Ligenfelter C, Vuong TT;

DR WPI; 1998-193616/17.

DR P-PSDB; AAW37883.

PT Breast cancer antigen 1 modulator protein - useful for diagnosing
 PT diseases involving unwanted cell growth, e.g. breast cancer, and for
 PT producing therapeutics for treatment of such diseases.

PS Claim 5; Fig 3; 73pp; English.

CC This cDNA clone, designated 091-132Q20 (ATCC 98143), which is not full-
 CC length, codes for a 46 kDa BRCA1 modulator protein (see AAW37883) that
 CC binds to the tumour suppressor gene product BRCA1, and which is
 CC characterised by a single leucine zipper domain. 3 cDNA clones (see also
 CC AAV29062-63) coding for BRCA1 modulator proteins (see AAW37881-83) were
 CC isolated from a HeLa cDNA library using a yeast two-hybrid assay with a
 CC GAL4-BRCA1(8-1293) fusion as bait. Vectors and host cells comprising the
 CC isolated nucleic acid sequences are claimed, as well as a process for
 CC producing BRCA1 modulator protein by culturing these host cells. BRCA1
 CC modulator proteins and nucleic acids can be used to diagnose diseases
 CC involving unwanted cell growth, e.g. breast cancer, and to identify
 CC compounds that alter BRCA1 interaction with BRCA1 modulators for the
 CC treatment of such diseases

SQ Sequence 1191 BP; 449 A; 228 C; 302 G; 212 T; 0 U; 0 Other;

Query Match

Best Local Similarity 1.2%; Score 20; DB 2; Length 1191;

Matches 20; Conservative 100.0%; Pred. No. 65;

Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCCAGAGGAGAGCTGGC 569

Db 451 ACCCCAGAGGAGAGCTGGC 470

RESULT 78

AAX86756
 ID AAX86756 standard; cDNA; 1191 BP.
 XX
 AC AAX86756;
 XX
 DT 27-OCT-1999 (first entry)
 XX
 DE cDNA 091-132Q20 encoding a BRCA1 modulator protein.
 XX
 KW Modulator protein; BRCA1; tumour suppressor protein; breast cancer;
 KW ovarian cancer; cell growth; cell proliferation; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 34..1191
 FT /*tag= a
 FT /note= "no termination codon given"
 XX
 PN US5948643-A.
 XX
 PD 07-SEP-1999.
 XX
 PF 13-AUG-1997; 97US-00968751.
 XX
 PR 13-AUG-1997; 97US-00968751.
 XX
 PA (ONIX-) ONYX PHARM INC.
 XX
 PI Rubinfeld B, Lingenfelter C, Vuong TT, Polakis PG;
 XX
 DR WPI; 1999-517952/43.
 DR P-PSDB; AAY30151.
 XX
 XX Modulator proteins that bind to and modulate the activity of the BRCA1
 PT tumor suppressor gene product, useful for the treatment of ovarian and
 PT breast cancer.
 XX
 PS Claim 1; Fig 3; 35pp; English.
 XX
 CC The present sequence encodes a modulator protein, that binds to and
 CC modulate the activity of the BRCA1 gene product (BRCA1). The BRCA1
 CC protein has been characterized as a tumour suppressor protein.
 CC Alterations in the amino acid sequence of BRCA1 causes breast and ovarian
 CC cancers by removing the controls on cell growth and proliferation.
 CC Research has shown that different regions on the BRCA1 molecule have
 CC different effects on cell growth and tumour suppression (e.g. full length
 CC truncated BRCA1 has no effect on breast cancer cell growth but will
 CC inhibit ovarian cancer cell growth). It has been suggested that different
 CC host cell factors (e.g. proteins) interact with different regions of the
 CC BRCA1 to control its function. The identification of these proteins (e.g.
 CC BRCA1MP) will facilitate the development of novel diagnostic methods and
 CC new therapeutics for identifying and treating cancers caused by changes
 CC in the expression or activity of BRCA1
 XX
 SQ Sequence 1191 BP; 449 A; 228 C; 302 G; 212 T; 0 U; 0 Other;
 Query Match 1.2%; Score 20; DB 2; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 550 ACCCCAGAGGAGCTGGC 569
 Db 451 ACCCCAGAGGAGCTGGC 470
 RESULT 79
 ACD13396
 ID ACD13396 standard; cDNA; 1618 BP.
 XX
 AC ACD13396;
 XX
 DT 13-AUG-2003 (first entry)
 XX
 DE Human DNA encoding a p53 modifier, SEQ ID 67.
 XX
 KW Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
 KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
 KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KW apoptotic disorder; cell proliferation disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200299122-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 03-JUN-2002; 2002WO-US017382.
 XX
 PR 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX
 DR WPI; 2003-156859/15.
 DR P-PSDB; ABO07222.
 XX
 XX Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.
 XX
 PS Example 2; Page 319-320; 678pp; English.
 XX
 CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway with a candidate modulator
 CC contacting a cell defective in p53 function with an HM amino acid
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored, modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence is an HM
 CC nucleic acid encoding a p53 pathway modifying protein
 XX
 SQ Sequence 1618 BP; 558 A; 330 C; 363 G; 367 T; 0 U; 0 Other;
 Query Match 1.2%; Score 20; DB 7; Length 1618;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 550 ACCCCAGAGGAGCTGGC 569
 Db 75 ACCCCAGAGGAGCTGGC 94

RESULT 80	
AAAS71711	
ID	AAAS71711 standard; cDNA; 1845 BP.
XX	
AC	AAAS71711;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #7515.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG07524.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 7515; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene map
CC	and in recombinant production of (II). The polynucleotides are also
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitat
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical ima
CC	of sites expressing (II). (I) and (II) are useful for treating disc
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiver
CC	and to produce other types of data and products dependent on DNA an
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagn
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SS	Sequence 1845 BP: 678 A: 424 C: 399 G: 344 T: 0 U: 0 Other:

Query Match	1.2%	Score 20;	DB 5;	Length 1845;
Best Local Similarity	100.0%;	Fred. No. 65;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1511	GGCCAGAGGGGCCCTGGGGCT	1530	
Db	92	GGCCAGAGGGGCCCTGGGGCT	111	

RESULT 81
AAH15042
ID AAH15042 standard; cDNA; 2045 BP.
XX AC
XX AAH15042;
DT XX
DT DT (first entry)
DE Human cDNA sequence SEQ ID NO:13025.
XX DE
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
OS Homo sapiens.
PN EP1074617-A2.
XX PD
PD 07-FEB-2001.
XX PF
PF 29-JUL-2000; 2000EP-00116126.
XX PR
PR 29-JUL-1999; 99JP-00248036.
XX PR
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
XX PR
PR 09-JUN-2000; 2000JP-00241899.
XX PA
PA (HELI-) HELIX RES INST.
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WFI; 2001-318749/34.
DR XX
PT PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX PS
PS Claim 8; SEQ ID NO 13025; 2537pp + Sequence Listing; English.
XX CC
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX SQ
SQ Sequence 2045 BP; 672 A; 421 C; 506 G; 446 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 4; Length 2045;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 550 ACCCAGGAGGAAGCTGGC 569
Db 502 ACCCAGGAGGAAGCTGGC 521

RESULT 82
 ID ABK70301 standard; cDNA; 2045 BP.
 AC ABK70301;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human lung cancer associated full length cDNA DMSM-181.
 XX
 KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200224057-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 20-SEP-2001; 2001WO-US042232.
 XX
 PR 22-SEP-2000; 2000US-0234837P.
 XX
 PR 10-OCT-2000; 2000US-0239440P.
 XX
 PR 29-JUN-2001; 2001US-0301928P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Benson DR, Mohamath R, Lodes MJ;
 XX
 DR WPI; 2002-372001/40.
 XX
 PT New tumor lung proteins and nucleic acids encoding the proteins, useful
 PT as vaccines and for treating, preventing, diagnosing or monitoring lung
 PT cancer.
 XX
 PS Claim 1; Page 175-176; 189pp; English.
 XX
 CC The invention relates to an isolated polynucleotide comprising a sequence
 CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
 CC or their fragments, homologues, variants or complements and their encoded
 CC polypeptides. Also included are an expression vector comprising the
 CC polynucleotide operably linked to an expression control sequence; a host
 CC cell transformed or transfected with an expression vector of; an isolated
 CC antibody, or its antigen-binding fragment that specifically binds to the
 CC polypeptide; a method for detecting the presence of a cancer in a patient
 CC ; a fusion protein comprising at least the polypeptide; an
 CC oligonucleotide that hybridizes to the polynucleotide under moderately
 CC stringent conditions; a method for stimulating and/or expanding T cells
 CC specific for a tumour protein; an isolated T cell population comprising T
 CC cells prepared from the method of above; a composition comprising a first
 CC component consisting of carriers and immunostimulants, and a second
 CC component selected from the polynucleotides, proteins, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells expressing the
 CC polypeptide; methods for stimulating an immune response or treating the
 CC cancer in a patient by administering the composition and diagnostic kits
 CC comprising at least one of the oligonucleotide of, or an antibody and a
 CC detection reagent consisting of a reporter group. The polypeptides and
 CC polynucleotides are useful as vaccines for the treatment or prevention of
 CC lung cancer, and for diagnosis and monitoring of such cancer. The
 CC polynucleotide, polypeptide and antigen presenting cells can be used to
 CC stimulate or expand T cells specific for a tumorous protein. The
 CC polynucleotides may be used as probes or primers for nucleic acid
 CC hybridisation, and in the preparation of ribozyme molecules for
 CC inhibiting expression of tumour polypeptides and proteins in tumour
 CC cells. The present sequence is one of the 183 lung cancer associated
 CC polynucleotides
 XX
 SQ Sequence 2045 BP; 672 A; 421 C; 506 G; 446 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 6; Length 2045;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCCAGAGGAGAGCTGGC 569
 DB 502 ACCCCAGAGGAGAGCTGGC 521
 RESULT 83
 AAX80156
 ID AAX80156 standard; cDNA; 4723 BP.
 XX
 AC AAX80156;
 XX
 DT 16-AUG-1999 (first entry)
 XX
 DE Neurite extending activity protein encoding cDNA.
 XX
 KW Neurite extending activity; anti-dementia; memory; brain function;
 KW dementia; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 196..2226
 XX /*tag= a
 PN JP11147897-A.
 XX
 PD 02-JUN-1999.
 XX
 PF 13-NOV-1997; 97JP-00331242.
 XX
 PR 13-NOV-1997; 97JP-00331242.
 XX
 PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 XX
 DR WPI; 1999-379889/32.
 DR P-PSDB; AAY17863.
 XX
 PT New protein with neurite extending activity - useful for treating
 PT dementia.
 XX
 PS Claim 3; Page 5-8; 9pp; Japanese.
 XX
 CC The present sequence encodes a protein (I) which has neurite extending
 CC activity. (I) may be administered to patients to prevent dementia or to
 CC improve memory and brain function. (I) is also used as a neurite-
 CC extending agent
 XX
 SQ Sequence 4723 BP; 1510 A; 1038 C; 1025 G; 1150 T; 0 U; 0 Other;
 Query Match 1.2%; Score 20; DB 2; Length 4723;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGCAGCAAGCTTGAGAGCCT 731
 DB 820 CGCAGCAAGCTTGAGAGCCT 839
 RESULT 84
 ABL08443
 ID ABL08443 standard; cDNA; 5736 BP.
 XX
 AC ABL08443;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 19811.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 XX

PR 14-AUG-2000; 2000US-0225366P.
PR 14-AUG-2000; 2000US-0225367P.
PR 14-AUG-2000; 2000US-0225368P.
PR 14-AUG-2000; 2000US-0225370P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 14-AUG-2000; 2000US-0226581P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-565190/63.

Nucleic acid encoding novel connective tissue associated polypeptides, used in diagnosing, preventing, treating or ameliorating a disorder such as cancer or rheumatoid arthritis.

Disclosure; SEQ ID NO 1624; 673pp; English.

The present invention relates to the isolation of novel human connective tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide (cDNA and genomic) sequences encoding them. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of diseases associated with connective tissue(s), including cancer. The polynucleotide sequences of the invention are also useful in gene therapy. ABK42102-ABK43116 represent genomic sequences encoding the novel human connective tissue related polypeptides. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 20029 BP; 5423 A; 4311 C; 4972 G; 5323 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 4; Length 20029;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1170 CACAGAGAGTTTGAGGAGT 1189

[illegible]

PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251388P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259578P.
 PR 17-JAN-2001; 2001US-00764847.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-634869/60.
 DR P-PSDB; ADB59617.
 XX
 PT New connective tissue-related polypeptides and polynucleotides, useful
 PT for treating, preventing and/or prognosing e.g. disorders of connective
 PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
 PT neoplasias.
 XX
 PS Disclosure; SEQ ID NO 1624; 248pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule (I), which
 CC comprises a sequence that is at least 95 % identical to a connective
 CC tissue-related polynucleotide encoding connective tissue antigens (CTA).
 CC The polypeptide or polynucleotide is useful for preventing, treating, or
 CC ameliorating medical conditions in a mammal. The connective tissue
 CC polypeptides, polynucleotides and antibodies are particularly useful for
 CC treating, preventing and/or prognosing disorders of connective tissues
 CC (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
 CC scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
 CC neoplasias (e.g. leukemia), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
 CC (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
 CC complications), autoimmune diseases (e.g. systemic lupus erythematosus,
 CC rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.
 CC

Query Match 1.2%; Score 20; DB 8; Length 20029;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1170 CACAGAGAAGTTTGAGGAGT 1189
 |||||
 Db 15291 CACAGAGAAGTTTGAGGAGT 15310

RESULT 88
 AAT61009
 ID AAT61009 standard; DNA; 39 BP.

XX AAT61009;
 AC
 XX 17-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 29-OCT-1997 (first entry)
 XX
 DE L-galactono-1,4-lacton:ferricytochrome-c-oxidoreductase DNA fragment 12.
 XX
 KW L-galactono-gamma-lactone; ascorbic acid; cauliflower; mitochondrion;
 KW antioxidant; chiral; L-galactonate; D-galacturonic acid; ss.
 XX
 OS Brassica oleracea var. botrytis.
 XX
 FH Key Location/Qualifiers
 CDS 1..39
 FT /*tag= a
 FT /product= "peptide fragment"
 FT /note= "Partial coding fragment, derived by reverse
 FT translation from directly sequenced peptide fragment of L
 FT -galactono-1,4-lacton: ferricytochrome-c-oxidoreductase,
 FT so does not include a termination codon"
 XX
 XX WO9704100-A2.
 PN
 PD 06-FEB-1997.
 XX
 PF 17-JUL-1996; 96WO-DE001339.
 XX
 PR 17-JUL-1995; 95DE-01025991.
 XX
 XX (FRAU) BRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
 XX
 XX Wissler J;
 XX
 DR WPI; 1997-132640/12.
 DR P-PSDB; AAW10612.
 XX
 PT Enzyme for oxidn. of L-galactono-gamma-lactone to ascorbic acid - useful
 PT for ascorbic acid synthesis from waste materials and determn. of chiral L
 PT -galactonate or D-galacturonic acid.
 XX
 PS Claim 6; Page 28; 34pp; German.
 XX
 CC A new L-galactono-1,4-lacton:ferricytochrome-c-oxidoreductase enzyme was
 CC isolated from cauliflower mitochondria. The enzyme oxidises L-galactono-
 CC gamma-lactone to ascorbic acid, has an apparent molecular weight of 56 kD
 CC and includes the amino acid sequence: Val-Gln-Gln-Leu-Val-Asp-Ala-Ile-Gln
 CC -Glu-Tyr-Gly-Leu. The present DNA sequence was derived by reverse
 CC translation from this specified fragment. A further 15 partial fragments
 CC of the enzyme have been obtained and sequenced. The enzyme is used for
 CC ascorbic acid synthesis, particularly from waste materials, and for the
 CC determination of chiral L-galactonate or D-galacturonic acid. Ascorbic
 CC acid is well known as a nutritional supplement, free radical scavenger
 CC and food antioxidant. The DNA molecule is used for the analysis and
 CC synthesis of equivalent structures, and associated functions. (Updated on
 CC 25-MAR-2003 to correct PR field.) (Updated on 17-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 39 BP; 7 A; 9 C; 15 G; 8 T; 0 U; 0 Other;

Query Match 1.2%; Score 19; DB 2; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1000 CAGCAGCTGCTGGATGCCA 1018
 |||||
 Db 4 CAGCAGCTGCTGGATGCCA 22

RESULT 89
 ABV76635
 ID ABV76635 standard; DNA; 41 BP.

XX AC ABV76635;
XX DT 21-FEB-2003 (first entry)
XX DE Human nuclear transcription factor 3 protein 21.56 probe 2.
XX KW Human; nuclear transcription factor 3 protein 21.56;
XX KW recombinant production; gene therapy; malignant tumour; cancer;
XX KW blood disease; HIV infection; human immunodeficiency virus;
XX KW immune disorder; inflammatory condition; cytostatic; antiinflammatory;
XX KW immunomodulator; probe; ss.
XX OS Homo sapiens.
XX PN CN1358764-A.
XX PD 17-JUL-2002.
XX PF 13-DEC-2000; 2000CN-00127895.
XX PR 13-DEC-2000; 2000CN-00127895.
XX PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX PI Mao Y, Xie Y;
XX WPI; 2002-733553/80.
XX DT Novel polypeptide-human nuclear transcription factor 3 protein 21.56 and
XX PT polynucleotide for encoding the polypeptide.
XX PS Example 6; Page 21 (Disclosure); 34pp; Chinese.
XX CC The invention relates to human nuclear transcription factor 3 protein
XX CC 21.56 (AB99971) and nucleic acids encoding it (ABV76629). The protein
XX CC has a molecular weight of 21.56 kD. The invention also relates to a
XX CC method for the recombinant production of the protein, an antagonist of
XX CC the protein, and the use of the protein, gene and antagonist in
XX CC therapeutic applications. Nuclear transcription factor 3 protein 21.56
XX CC can be used in the treatment of a variety of diseases such as malignant
XX CC tumours, blood diseases, HIV (human immunodeficiency virus) infection,
XX CC immune disorders and inflammatory conditions. Sequences ABV76634-ABV76635
XX CC represent human nuclear transcription factor 3 protein 21.56 probes used
XX CC in an exemplification of the invention
XX SQ Sequence 41 BP; 15 A; 7 C; 12 G; 7 T; 0 U; 0 Other;
Query Match 1.2%; Score 19; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 615 TTCACAGACGATGAG 633
DB 5 TTCACAGACGATGAG 23
RESULT 90
AAL24011
ID AAL24011 standard; cDNA; 105 BP.
XX AC AAL24011;
XX DT 07-DEC-2001 (first entry)
XX DE Human breast cancer expressed polynucleotide 16468.
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX OS Homo sapiens.
XX PN WO200151628-A2.
XX XX

PD 19-JUL-2001.
XX 10-JAN-2001; 2001WO-US000798.
XX 14-JAN-2000; 2000US-0176077P.
XX PR 14-MAR-2000; 2000US-0189167P.
XX PR 24-MAR-2000; 2000US-0192099P.
XX PR 29-MAR-2000; 2000US-0193480P.
XX PR 15-MAY-2000; 2000US-0205230P.
XX PR 09-JUN-2000; 2000US-0211315P.
XX PR 25-JUL-2000; 2000US-0220534P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX DT New peptide useful as a marker for the diagnosis of breast cancer.
XX PT Claim 1; Page 3014; 3695pp; English.
XX PS The invention relates to human breast cancer expressed polynucleotides
XX CC (AAL07544-AAL26789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, monitoring, characterising treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX CC activity
XX SQ Sequence 105 BP; 23 A; 31 C; 39 G; 12 T; 0 U; 0 Other;
Query Match 1.2%; Score 19; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 AGAAGTCCCGGCGACGAGC 147
DB 1 AGAAGTCCCGGCGACGAGC 19
RESULT 91
AAL15165
ID AAL15165 standard; cDNA; 109 BP.
XX AC AAL15165;
XX DT 07-DEC-2001 (first entry)
XX DE Human breast cancer expressed polynucleotide 7622.
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX OS Homo sapiens.
XX PN WO200151628-A2.
XX PD 19-JUL-2001.
XX PF 10-JAN-2001; 2001WO-US000798.
XX PR 14-JAN-2000; 2000US-0176077P.
XX PR 14-MAR-2000; 2000US-0189167P.
XX PR 24-MAR-2000; 2000US-0192099P.
XX PR 29-MAR-2000; 2000US-0193480P.
XX PR 15-MAY-2000; 2000US-0205230P.
XX PR 09-JUN-2000; 2000US-0211315P.
XX PR 25-JUL-2000; 2000US-0220534P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX XX

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 1371; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides (AA107544-AA126789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

XX SQ Sequence.109 BP; 25 A; 32 C; 38 G; 14 T; 0 U; 0 Other;

Query Match 1.2%; Score 19; DB 4; Length 109;

Best Local Similarity 100.0%; Pred. NO. 28+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 AGAAGGTCCCGGAGCAGC 147

DB 1 AGAAGGTCCCGGAGCAGC 19

RESULT 92

ABT21876/c

ID ABT21876 standard; DNA; 138 BP.

XX AC ABT21876;

XX DT 16-APR-2003 (first entry)

XX DE Breast cancer marker gene SEQ ID No 249.

XX KW Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen; chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid; surrogate marker gene; pharmacodynamic marker gene; transgenic animal; human; ds.

XX OS Homo sapiens.

XX PN WO200285298-A2.

XX PD 31-OCT-2002.

XX PF 19-APR-2002; 2002WO-US012612.

XX PR 20-APR-2001; 2001US-0285163P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;

XX WPI; 2003-093053/08.

XX Novel isolated polypeptide encoded by breast cancer marker gene, useful for diagnosing, staging, monitoring, prognosing and treating diseases associated with breast cancer.

XX Disclosure; Page 126; 725pp; English.

XX The invention relates to an isolated polypeptide encoded by a breast cancer marker gene comprising any of 1417, 21-805 nucleotide sequences, given in the specification. The methods of the invention are useful for diagnosing patients having an identified breast mass or symptoms associated with breast cancer, to diagnose breast cancer or its precursors, and for monitoring the efficacy of treatment of a breast cancer patient (e.g. efficacy of chemotherapy). The methods are also

CC useful for evaluating a patient before, after or during therapy, to evaluate the reduction in a tumour burden. The breast cancer marker gene proteins are useful as immunogens for raising antibodies, by immunising a mammal with a breast cancer marker protein. The marker proteins are useful as bait proteins in a two-hybrid or three-hybrid assay, to identify other proteins which bind to or interact with the marker proteins. The breast cancer marker genes are useful as surrogate marker genes for one or more disorders, disease states or conditions leading to disease states, in particular, breast cancers. The breast cancer marker genes are useful as pharmacodynamic marker genes. An antibody which selectively binds to a protein of a breast cancer marker gene is useful for treating cancers, particularly breast cancers. The host cell of the invention is useful for producing non-human transgenic animals. This polynucleotide sequence represents one of the breast cancer marker genes of the invention

XX SQ Sequence 138 BP; 16 A; 48 C; 46 G; 28 T; 0 U; 0 Other;

Query Match 1.2%; Score 19; DB 7; Length 138;

Best Local Similarity 100.0%; Pred. NO. 28+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 AGAAGGTCCCGGAGCAGC 147

DB 138 AGAAGGTCCCGGAGCAGC 120

RESULT 93

ABV14515/c

ID ABV14515 standard; cDNA; 288 BP.

XX AC ABV14515;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 14506.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 2422; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in tables 1-9 (ABV0010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 288 BP; 52 A; 78 C; 83 G; 75 T; 0 U; 0 Other;

Query Match 1.2%; Score 19; DB 5; Length 288;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 AGAAGTCCCGCAGCAGC 147
 DB 288 AGAAGTCCCGCAGCAGC 270

RESULT 94
 AAC02994/c
 ID RAC02994 standard; cDNA; 396 BP.
 XX
 AC AAC02994;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 2992.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GEST) GENSET.
 XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR P-PSDB; AAG02988.
 DR
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 2992; 71pp + Sequence Listing; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors
 XX
 SQ Sequence 396 BP; 85 A; 128 C; 83 G; 96 T; 0 U; 4 Other;

Query Match 1.2%; Score 19; DB 3; Length 396;
 Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1235 AGCAGGAGTGGAAAGAT 1253
 DB 19 AGCAGGAGTGGAAAGAT 1
 RESULT 95
 ABN62601/c
 ID ABN62601 standard; cDNA; 550 BP.
 XX
 AC ABN62601;
 XX
 DT 28-JUN-2002 (first entry)
 XX
 DE Human cancer related polynucleotide SEQ ID NO 2568.
 XX
 KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
 KW gene therapy; cancer; tumour; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200214500-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US025840.
 XX
 PR 16-AUG-2000; 2000US-0236326P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
 PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
 XX
 DR WPI; 2002-241905/29.
 XX
 PT New nucleic acid for producing a polypeptide, detecting differentially
 PT expressed genes correlated with a cancerous state of a mammalian cell,
 PT and inhibiting tumor growth.

PS Claim 1; SEQ ID NO 2568; 893pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
 CC with cytostatic activity. The polynucleotide is used to produce a
 CC polypeptide, to detect differentially expressed genes correlated with a
 CC cancerous state of a mammalian cell and to inhibit tumour growth. The
 CC polynucleotide is used as a probe in mapping and tissue profiling. The
 CC encoded polypeptide and antibodies to the polypeptide can also be used
 CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
 CC gene therapy. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 550 BP; 127 A; 108 C; 155 G; 158 T; 0 U; 2 Other;

Query Match 1.2%; Score 19; DB 6; Length 550;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CAAAAGAACGGGGTGTCC 33
 DB 60 CAAAAGAACGGGGTGTCC 42

RESULT 96
 ABN62686/c
 ID ABN62686 standard; cDNA; 592 BP.
 XX
 AC ABN62686;
 XX
 DT 28-JUN-2002 (first entry)

XX DE Human cancer related polynucleotide SEQ ID NO 2653.
XX KW Human; cytosolic; gene expression; gene mapping; tissue profiling;
XX KW gene therapy; cancer; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200214500-A2.
XX PD 21-FEB-2002.
XX PF 16-AUG-2001; 2001WO-US025840.
XX PR 16-AUG-2000; 2000US-0226326P.
XX PA (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
XX PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX DR WPI; 2002-241905/29.
XX XX New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX XX Claim 1; SEQ ID NO 2653; 883pp + Sequence Listing; English.
XX CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 592 BP; 138 A; 125 C; 161 G; 168 T; 0 U; 0 Other;
Query Match 1.2%; Score 19; DB 6; Length 592;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 CAAAAAGACGGGGCTGCC 33
Db 60 CAAAAAGACGGGGCTGCC 42
RESULT 97
ABK77260/c
ID ABK77260 standard; DNA; 627 BP.
XX AC
XX AC ABK77260;
XX DT 13-AUG-2002 (first entry)
XX DE Bacillus clausii genomic sequence tag (GST) #103.
XX KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX OS Bacillus clausii.
XX PN WO200229113-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-US031437.

XX PR 06-OCT-2000; 2000US-00680598.
XX PR 27-MAR-2001; 2001US-0279526P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
XX PA (NOVO) NOVOZYMES AS.
XX PI Berka R, Clausen IG;
XX PI WPI; 2002-416684/44.
XX DR Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX PS Claim 11; SEQ ID NO 4551; 200pp; English.
XX CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 627 BP; 170 A; 149 C; 174 G; 134 T; 0 U; 0 Other;
Query Match 1.2%; Score 19; DB 6; Length 627;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GTCTTCAACACAGGACC 994
Db 452 GTCTTCAACACAGGACC 434
RESULT 98
AAH01784
ID AAH01784 standard; DNA; 668 BP.
XX AC
XX AC AAH01784;
XX DT 24-JUL-2001 (first entry)
XX DE Klebsiella oxytoca nucleotide sequence SEQ ID NO:1777.
XX KW Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitological;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial; vaccine;
KW primer; ds.
XX OS Klebsiella oxytoca.
XX PN WO200123604-A2.
XX PF 05-APR-2001.

[illegible]

```

OS Humicola insolens.
XX
XX Key Location/Qualifiers
FH CDS 43..783
FT /*tag= a
FT sig_peptide 43..147
FT /*tag= b
FT mat_peptide 148..780
FT /*tag= c
XX
XX WO9613580-A1.
XX
XX 09-MAY-1996.
XX
XX 26-OCT-1995; 95WO-DK000427.
XX
XX 26-OCT-1994; 94DK-00001240.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Sandal T.; Kauppinen S.; Kofod LV;
XX
XX WPI; 1996-239493/24.
XX P-FSDB; AAR95051.
XX
XX Isolated DNA encoding lipolytic enzyme - derived from Humicola insolens
XX DSM1800, used partic. as detergent additive for laundry and dishwashing
XX compsns.
XX
XX Claim 1; Page 51; 63pp; English.
XX
XX The present sequence encodes an enzyme with lipolytic activity, derived
XX from Humicola insolens DSM1800. The enzyme has a mol. wt. of about 20-21
XX kDa, a pI in the range of 709, a pH optimum of about 8 and has
XX specificity towards short-chain lipid substrates. It can be used as a
XX detergent additive, partic. in compsns. for laundry washing or
XX dishwashing
XX
XX Sequence 1016 BP; 231 A; 319 C; 289 G; 177 T; 0 U; 0 Other;
XX
XX Query Match 1.2%; Score 19; DB 2; Length 1016;
XX Best Local Similarity 100.0%; Pred. No. 2e+02;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1307 GGGAGAGCAGCAACAGGC 1325
XX |||||
XX Db 95 GGGAGAGCAGCAACAGGC 77

```

Search completed: June 8, 2004, 13:50:04
 Job time : 610.16 secs

C	66	21	1.3	608	6	AX256016	AX256016 Sequence
C	67	21	1.3	833	9	HUMCYCLO	M60457 Human cyclo
C	68	21	1.3	851	6	AR274927	AR274927 Sequence
C	69	21	1.3	851	9	HUMCYBPA	M60857 Human cyclo
C	70	21	1.3	893	6	AX409730	AX409730 Sequence
C	71	21	1.3	893	6	AX779845	AX779845 Sequence
C	72	21	1.3	893	6	AX779846	AX779846 Sequence
C	73	21	1.3	893	9	HUMSCYLP	M63573 Human scire
C	74	21	1.3	894	9	BC001125	BC001125 Homo sapi
C	75	21	1.3	897	9	BC032138	BC032138 Homo sapi
C	76	21	1.3	908	9	BC020800	BC020800 Homo sapi
C	77	21	1.3	1019	9	BC008848	BC008848 Homo sapi
C	78	21	1.3	1060	3	AY191780	AY191780 Branchios
C	79	21	1.3	2739	9	PP1GACR	X53704 P.pygmaeus
C	80	21	1.3	8821	9	HSW804617	AL933304 Homo sapi
C	81	21	1.3	9708	9	AF528039	AF528039 Homo sapi
C	82	21	1.3	113802	9	AL137003	AL137003 Human DNA
C	83	21	1.3	128422	9	AL357139	AL357139 Human DNA
C	84	21	1.3	128461	9	AL390741	AL390741 Human DNA
C	85	21	1.3	147124	2	AC022768	AC022768 Homo sapi
C	86	21	1.3	157032	2	AC118107	AC118107 Rattus no
C	87	21	1.3	158801	9	AC079118	AC079118 Homo sapi
C	88	21	1.3	159281	2	AC134703	AC134703 Rattus no
C	89	21	1.3	161728	2	AC108791	AC108791 Mus muscu
C	90	21	1.3	164857	9	AC110054	AC110054 Homo sapi
C	91	21	1.3	169942	9	AC124793	AC124793 Homo sapi
C	92	21	1.3	173043	2	AC124792	AC124792 Homo sapi
C	93	21	1.3	1734287	9	AC100840	AC100840 Homo sapi
C	94	21	1.3	174677	9	AC063960	AC063960 Homo sapi
C	95	21	1.3	175258	2	AC017098	AC017098 Homo sapi
C	96	21	1.3	178344	2	AC067846	AC067846 Homo sapi
C	97	21	1.3	178611	2	AC024631	AC024631 Homo sapi
C	98	21	1.3	183102	2	AC146968	AC146968 Sus scrof
C	99	21	1.3	211601	2	AC146969	AC146969 Sus scrof
C	100	21	1.3	212026	2	AC106432	AC106432 Rattus no

ALIGNMENTS

```

RESULT 1
AR409338 LOCUS linear PAT 18-DEC-2003
DEFINITION 1638 bp DNA
Sequence 46 from patent US 6632923.
AR409338 ACCESSION
AR409338.1 GI:40160126
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
Unclassified.
1 (bases 1 to 1638)
Lee,S.A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
Patent: US 6632923-A 46 14-OCT-2003;
Location/Qualifiers
1..1638
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1638; DB 6; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAACCGAGACAAAAGAACGGGGCTGCCAAACAATCCAAATCCAAAGACGACCCA 60
DB 1 ATGAAGAACCAAGACAAAAGAACGGGGCTGCCAAACAATCCAAATCCAAAGACGACCCA 60

QY 61 GGACAAACCGGAAGACGAGACCCGAGGAGAGCCAGGAGCGGCCCCAGCCAGCGGCTCCTGCA 120
DB 61 GGACAAACCGGAAGACGAGACCCGAGGAGAGCCAGGAGCGGCCCCAGCCAGCGGCTCCTGCA 120

```



```
Db 1201 CTTTCCAAAGCAGGAGGTATTCCACCATTTCAAGCAGGAGATGGAAGAATGACTAAG 1260
Qy 1261 AAGATCAAGAGAGCTGGAGAAAGAAACACCATGTATCCCGTCCCGTGGGAGAGAGCAAC 1320
Db 1261 AAGATCAAGAGAGCTGGAGAAAGAAACACCATGTATCCCGTCCCGTGGGAGAGAGCAAC 1320
Qy 1321 AAGGCCCTGCTTGTAGATGGCTGAGAGAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1380
Db 1321 AAGGCCCTGCTTGTAGATGGCTGAGAGAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1380
Qy 1381 CAGGTAAATAATCCACGGCTGGAGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1440
Db 1381 CAGGTAAATAATCCACGGCTGGAGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1440
Qy 1441 CTGAACAAGAGAGGTACAGCACTGTAGTCTGTGTGCGGAGGCTCCCTCACTGACAGTGGC 1500
Db 1441 CTGAACAAGAGAGGTACAGCACTGTAGTCTGTGTGCGGAGGCTCCCTCACTGACAGTGGC 1500
Qy 1501 CTTGAGAGAGGAGGAGGAGGCTGGGCTCAAGACACCCAGCTCCCGCAGGCTCACAGAA 1560
Db 1501 CTTGAGAGAGGAGGAGGAGGCTGGGCTCAAGACACCCAGCTCCCGCAGGCTCACAGAA 1560
Qy 1561 GCGCCTTGTCTACCCAGGAGCACCAGACAGAGCATCAGGCCAGACTGGGCTCAAGAG 1620
Db 1561 GCGCCTTGTCTACCCAGGAGCACCAGACAGAGCATCAGGCCAGACTGGGCTCAAGAG 1620
Qy 1621 CCCACCTCGCCAGAGGCC 1638
Db 1621 CCCACCTCGCCAGAGGCC 1638

RESULT 2
AX239603 LOCUS AX239603 1638 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 46 from Patent WO0164874.
ACCESSION AX239603
VERSION AX239603.1 GI:15797277
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
Patent: WO 0164874-A 46 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
source
1..1638
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..>1638
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC88300.1"
/db_xref="GI:15797278"
/db_xref="REFMTEMBL:CAC88300"
/translation="MKNDKNGAKQSNPKSSPGQPEAGPEGAQERPSQAPAVEAE
GPGSQAPRKPEGAQRTAQGALRDSEELSRQLEDILSTYCDVNNQGGPEGAGQ
EPAPEDAEKRTYVARGNEPEPTPVVNGKEPKSGDPNTEEIRQSDVEGDRHRPQ
EKKAAGLGEKRTITLMQTLNTLSTPEEKLAALCKRYALELHEHNSQKMKLLQKQS
OLVQKDLHGEHSKAVLARSKLRELOHNRSLKEGVQVAREEERKEKVTSH
FOVTLNDLOMEONHNSKLROENMELAEERKLKLIROYELREHIDKVPKHLQ
OLVAKLQQAQAEERHOREKDFLLKEAVSQRMCELMKQETHLKQALALYE
KFEQNTLSKSEVPTTFKQEMEKWTKIKLEKETMYRSRNSKLLPERRPQ
TVRDKLEGLQVKIQRLEKLCALQTERNDLNKRVDLSAGGGGSLTDSGPERPEP
GAQAPSPRVTEAPCPGAPSTEASGTPQEPSTARA"
ORIGIN
Query Match 100.0%; Score 1638; DB 6; Length 1638;
```

1021	CTTCCAGAGGCCACGAGGATGCTAAAGAGGCGCAGAAAGCGGCA	COACGCGGGAGAAAGGAT	1080
	DB		
1081	TTTCTCCTAAAGAGGCAGT	TAGAGTCCGAGAGGATGTGTGAGCTGATGAAGCAGCAAGAG	1140
	QY		
1081	TTTCTCCTGAAGAGGCAGT	TAGAGTCCGAGAGGATGTGTGAGCTGATGAAGCAGCAAGAG	1140
	DB		
1141	ACCCACTGAAGCAACAGCTT	TGCCCTATACACAGAGAAATTTGAGAGTTCAGAAACA	1200
	QY		
1141	ACCCACTGAAGCAACAGCTT	TGCCCTATACACAGAGAAATTTGAGAGTTCAGAAACA	1200
	DB		
1201	CTTTCCAAAGCAGCGAGTAT	TACCAACATTAACAGCAGGATGAAAAGATGACTAAG	1260
	QY		
1201	CTTTCCAAAGCAGCGAGTAT	TACCAACATTAACAGCAGGATGAAAAGATGACTAAG	1260
	DB		
1261	AAAGATCAAGAAGCTGAGAGAAAGAAAC	CACCATGTACCGGTCCGGTGGGAGAGCAGCAAC	1320
	QY		
1261	AAAGATCAAGAAGCTGAGAGAAAGAAAC	CACCATGTACCGGTCCGGTGGGAGAGCAGCAAC	1320
	DB		
1321	AAGGCCCTCTTGAGATGGCT	TAGGAGAAACAGTCCGGGATTAAGAACTGAGGGGCCCTG	1380
	QY		
1321	AAGGCCCTCTTGAGATGGCT	TAGGAGAAACAGTCCGGGATTAAGAACTGAGGGGCCCTG	1380
	DB		
1381	CAGGTAAAAATCCAACGGCTTGAGAGAGCTGTG	CCGGGCACTGCAGACAGAGCGCAATGAC	1440
	QY		
1381	CAGGTAAAAATCCAACGGCTTGAGAGAGCTGTG	CCGGGCACTGCAGACAGAGCGCAATGAC	1440
	DB		
1441	CTGAACAAGAGGGTACAGGACCTGAGTGTGTG	TGGCCAGGGGTCTCTACTGACAGTGGC	1500
	QY		
1441	CTGAACAAGAGGGTACAGGACCTGAGTGTGTG	TGGCCAGGGGTCTCTACTGACAGTGGC	1500
	DB		
1501	CCTGAGAGAGGCCAGAGGGGCTTGGGCTCAAGCA	CCCGAGTCCCCCAGGGTTCAGAA	1560
	QY		
1501	CCTGAGAGAGGCCAGAGGGGCTTGGGCTCAAGCA	CCCGAGTCCCCCAGGGTTCAGAA	1560
	DB		
1561	GGCCCTTGTTACCCAGAGCACCGAGCA	CAGAAAGCATCAGGCCAGACTGGGCCCTCAAGAG	1620
	QY		
1561	GGCCCTTGTTACCCAGAGCACCGAGCA	CAGAAAGCATCAGGCCAGACTGGGCCCTCAAGAG	1620
	DB		
1621	CCCACTCCGCCAGGGCC	1638	
	QY		
1621	CCCACTCCGCCAGGGCC	1638	
	DB		

```

RESULT 3
AF516206 LOCUS linear PRI 15-FEB-2003
AF516206 Homo sapiens taxilin mRNA, complete cds.
AF516206 AF516206.1 GI:28394223
AF516206.1 Homo sapiens (human)
AF516206 Homo sapiens
AF516206 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AF516206 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AF516206 1 (bases 1 to 1993)
AF516206 Nogi, S., Satoh, S., Nakano, M., Shimizu, H., Fukushima, H.,
AF516206 Mariyama, A., Terano, A. and Shirataki, H.
AF516206 Taxilin; a novel syntaxin-binding protein that is involved in
AF516206 Ca2+-dependent exocytosis in neuroendocrine cells
AF516206 Genes Cells 8 (1), 17-28 (2003)
AF516206 2 (bases 1 to 1993)
AF516206 Shirataki, H., Nogami, S. and Satoh, S.
AF516206 Direct Submission
AF516206 Submitted (29-MAY-2002) Division of Molecular and Cell Biology,
AF516206 Institute for Medical Sciences, Dokkyo University School of
AF516206 Medicine, 880 Kitakobayashi, Mibu-machi, Tochigi 321-0293, Japan
AF516206 Location/Qualifiers
AF516206 1. .1993
AF516206 /organism="Homo sapiens"
AF516206 /mol_type="mRNA"
AF516206 /db_xref="taxon:9606"
FEATURES
source

```

ORGANISM	Hom sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 4851) Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeinger, A., Pobo, G., Han, M. and Wiemann, S.
TITLE	Direct Submission
JOURNAL	Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BWFZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseidorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp451K215) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.
FEATURES	Location/Qualifiers 1..4851 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="RZPD:DKFZp451K215" /db_xref="taxon:9606" /map="1p34.2-36.11" /clone="DKFZp451K215" /tissue_type="human skeletal muscle" /clone_lib="451 (synonym: hiccl). Vector pSport1; host DH10B; sites NotI + SalI" /dev_stage="adult" 1..4851 /gene="DKFZp451K215" 103..1743 /gene="DKFZp451K215" /codon_start=1 /product="hypothetical protein" /protein_id="CAD91138.1" /db_xref="GI:30722346" /translation="MKNQDKKNGAAKQSNPKSSPQPEAGPEGHQPERSQAAPAVEAE GPGSQAPRRPEGAQARTQAQSGALRDSSELSRQLEDILSTYICVNNQGGEGDGAQ EPAPSDAKSRITYVARNGEPEPTVPVNGEKEPEPSKDNPTEIIRQSDVEVGDHRRQ EKKAAGKKEITLLMQTLNTLSTPEELKALCKYAELEHRSNOKMLLQKQK QLVQKXDLHRLGSHKAVLARSKLESRLQHNRLSKLEGVQVRAREEEKKRKVTH FTLNDIOLQEHNRNSKLQENBELARLKLIEYELREHIDKVPKHDLOQ QLVDKLAQQAQMLKEABERHOREXDFLLKEAVSORMCELMKQOETHLKOOLALVE KPEEPNTLSKSSEVPTFKQEMEXWKIKKLEKETTMYSRWESSNKALLEAREK TVADKELBGQVKIQRLEKLCALQTERNDLNKRVQDLSAGGQSLDSDSGERRPEP GAQAPSPPRVTEAPCYPGAPSAESAGTGPQEPSTARS" 4814..4819 /gene="DKFZp451K215" 4834 /gene="DKFZp451K215"
ORIGIN	Query Match 90.7%; Score 1485; DB 9; Length 4851; Best Local Similarity 99.9%; Pred. No. 0; Matches 1585; Conservative 0; Mismatches 2; Indels 0; Gaps 0; QY 1 ATGAAGAACCAACACAAAAGAACCGGGCTGCCAACCAATCAATCCAAAAGCAGGCCA 60 DB 103 ATGAAGAACCAACACAAAAGAACCGGGCTGCCAACCAATCAATCCAAAAGCAGGCCA 162 QY 61 GGACAAACCGGAGCAGGACCCGAGGAGAGCCAGGAGCGGCCAGCCAGCGCGCTCTGCA 120 DB 163 GGACAAACCGGAGCAGGACCCGAGGAGAGCCAGGAGCGGCCAGCCAGCGCGCTCTGCA 222 QY 121 GTAGAGCAGAGGTCCTCCGGCAGCAGCGGCTCTCGAAGCCGAGGCTCAAGCC 180 DB 223 GTAGAGCAGAGGTCCTCCGGCAGCAGCGGCTCTCGAAGCCGAGGCTCAAGCC 282 QY 181 AGAACCGCTCAGTCTCGGGCGCTTGTGATGTCTGTGAGGAGCTGAGCGCGCAACTGAA 240

Db 283 AGAACGGCTCAGTCTGGGCGCCTTCGTGATGTCTCTGAGGAGCTGAGCGCCCACTGGAA 342
QY 241 GACATCTGAGCACATACATCTGTGACCAATAACCAAGGGGGGGCCCCGGGCGAGGATGGGCA 300
Db 343 GACATCTGAGCACATACATCTGTGACCAATAACCAAGGGGGGGCCCCGGGCGAGGATGGGCA 402
QY 301 CAGGGTGAGCGGCTGAACCCGAGAGTGCAGAGAGTCCCGGACCTATGTGCGAAGGAAT 360
Db 403 CAGGGTGAGCGGCTGAACCCGAGAGTGCAGAGAGTCCCGGACCTATGTGCGAAGGAAT 462
QY 361 GGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAAAGAACCTCCCAAGGGGATCCA 420
Db 463 GGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAAAGAACCTCCCAAGGGGATCCA 522
QY 421 AACACAGAGAGATCCGGCAGAGTGCAGAGTGCAGAGCCGAGACCATCGAAGGCCACAG 480
Db 523 AACACAGAGAGATCCGGCAGAGTGCAGAGTGCAGAGCCGAGACCATCGAAGGCCACAG 582
QY 481 GAGAAGAAAAAGCCCAAGGGTTGGTAAAGAGATCACGTTGCTGATGCAGACATTGAAT 540
Db 583 GAGAAGAAAAAGCCCAAGGGTTGGGGAAGAGATCAAGTTGCTGATGCAGACATTGAAT 642
QY 541 ACTCTGAGTACCCAGAGAGAGCTGGCTGTCTGTGCAAGAGTATGCTGAAGTCTG 600
Db 643 ACTCTGAGTACCCAGAGAGAGCTGGCTGTCTGTGCAAGAGTATGCTGAAGTCTG 702
QY 601 GAGGAGCAGCGGAATTCACAGAGAGAGTGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
Db 703 GAGGAGCAGCGGAATTCACAGAGAGAGTGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 762
QY 661 GTCCAAAGAGAGACCACTCGCGGTGAGCAAGCAAGGCCGTCTGCGCCGCGAGCAAG 720
Db 763 GTCCAAAGAGAGACCACTCGCGGTGAGCAAGCAAGGCCGTCTGCGCCGCGAGCAAG 822
QY 721 CTTGAGAGCTATGCGGTGAGTGCAGCGGCAACCGCTCCTCAAGGAAGAGTGTG 780
Db 823 CTTGAGAGCTATGCGGTGAGTGCAGCGGCAACCGCTCCTCAAGGAAGAGTGTG 882
QY 781 CAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGGTGACCTCGCACTTCCAGTGTACA 840
Db 883 CAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGGTGACCTCGCACTTCCAGTGTACA 942
QY 841 CTGAATGACATTCAGCTGCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 943 CTGAATGACATTCAGCTGCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1002
QY 901 GAGAACATGAGTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGTGGCGGAG 960
Db 1003 GAGAACATGAGTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGTGGCGGAG 1062
QY 961 GAGCATATGACAAAGTCTTCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1063 GAGCATATGACAAAGTCTTCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
QY 1021 CTCAGCAGCGCCAGAGAGATGCTAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1123 CTCAGCAGCGCCAGAGAGATGCTAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1182
QY 1081 TTTTCTCTGAAAGAGGAGGAGTGTGAGGATGTTGAGTGTGATGAGGAGAGAGAGAG 1140
Db 1183 TTTTCTCTGAAAGAGGAGGAGTGTGAGGATGTTGAGTGTGATGAGGAGAGAGAGAG 1242
QY 1141 ACCCACTGAGCAACAGCTTCCCTATACAGAGAGAGTTCAGAGAGTTCAGAGACACA 1200
Db 1243 ACCCACTGAGCAACAGCTTCCCTATACAGAGAGAGTTCAGAGAGTTCAGAGACACA 1302
QY 1201 CTTTCCAAAAGCAGGAGTATTCCACCAATTCAGCAGAGAGTGAAGAAAGATGACTAAG 1260
Db 1303 CTTTCCAAAAGCAGGAGTATTCCACCAATTCAGCAGAGAGTGAAGAAAGATGACTAAG 1362
QY 1261 AAGATCAAGAGCTGGAG 1320
Db 1363 AAGATCAAGAGCTGGAG 1422

QY 1321 AAGGCCCTGCTTGTGAGTGGTGGAGAGAAAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1380
Db 1423 AAGGCCCTGCTTGTGAGTGGTGGAGAGAAAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1482
QY 1381 CAGGTAAAAATCCAAACGGCTGGAGAGTGTGCGGGGCACTGCAGACAGAGCGCAATGAC 1440
Db 1483 CAGGTAAAAATCCAAACGGCTGGAGAGTGTGCGGGGCACTGCAGACAGAGCGCAATGAC 1542
QY 1441 CTGAACAGAGAGGTACAGGACCTGAGTGTGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
Db 1543 CTGAACAGAGAGGTACAGGACCTGAGTGTGTGGCCAGGGCTCCCTCACTGACAGTGGC 1602
QY 1501 CCTGAGAGAGGCGCAGAGGCGCTGCGGCTCAAGACCCAGCTCCCGCAGGGTCACAGAA 1560
Db 1603 CCTGAGAGAGGCGCAGAGGCGCTGCGGCTCAAGACCCAGCTCCCGCAGGGTCACAGAA 1662
QY 1561 GCGCCTTGTCTACCCAGAGCAACCGAGC 1587
Db 1663 GCGCCTTGTCTACCCAGAGCAACCGAGC 1689

RESULT 5
HSM803946 4880 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKFp451I0918 (from clone DKFp451I0918);
DEFINITION complete cds.
ACCESSION AL832636
VERSION AL832636.2 GI:30268324
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4880)
AUTHORS Ansoorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
On Apr 30, 2003 this sequence version replaced gi:21733212.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFp451I0918) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
Location/Qualifiers
1..4880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFp451I0918"
/db_xref="taxon:9606"
/map="lp34.2-36.11"
/clone="DKFp451I0918"
/tissue_type="human spinal cord"
/clone_lib="451 (synonym: hicc1). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4880
/gene="DKFp451I0918"
125..1765
/gene="DKFp451I0918"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD89951.1"
/db_xref="GI:30268325"
/translation="MKQDKKNGAAQQSNPKSPGQPEAGPEGAQERFQAAPAVEAE"

GPCSQAOPKPEGAQARTASGALRDVSELSROLEDILITCYDNNQGGPGEDGAQG
EPAPEDAKSRTYVARNGEPEPTPVNGKEPSKGDPTNEIRDSDEVDGDRHRPQ
EKQAKDGLKIBLLMOTUNTISTPEEKLAALCKYAEELLEHRNSQOMKLQKQS
QLVQEKDHLRGSHSKAVLARSKLSELRQLNRSLKEEGVQRAREEERKEVTS
FQVTLNDIQLOMEQHNRNSKLROENMELAEKLKLI EOYELRESHIDKVFKHDLQO
QLVADKLOQAQMLKEABERHOREKDFLLKEAVEORMCELMKQOETHILKQOLALYE
KPEEFONTLSKSEVFTTPEOBEMKMTKKIKLEKETTWYRSKSSNKALLEMAEEK
TVRDXLEGLQVUKIOPLEKLCRALOTERNDLNKRVDLSAGGSGSLTDSGPPRRPEP
GQAQSSPRVTPACVPFGAPSTEAASQOTGPQPTBARA

polyA_signal

4833..4838

/gene="DKFZp451I0918"

polyA_site

4852

/gene="DKFZp451I0918"

ORIGIN

Query Match 88.6%; Score 1451; DB 9; Length 4880;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

35 AACATCCCAATCCAAAAGCAGCCAGGACAAACCGGAAGCAGACCCGAGGAGGCCAGG 94
159 AACATCCCAATCCAAAAGCAGCCAGGACAAACCGGAAGCAGACCCGAGGAGGCCAGG 218
95 AGCGGCCAGCAGCGGGTCTCTGAGTAGAGAGAGAGAGGTCCCGCAGCAGCCAGGCTC 154
219 AGCGGCCAGCAGCGGGTCTCTGAGTAGAGAGAGAGAGGTCCCGCAGCAGCCAGGCTC 278
155 CTCGGAAGCGGAGGGTGTCTAAGCCAGAACCGCTCAGTCTGGGCCCTTCTGTGATGCT 214
279 CTCGGAAGCGGAGGGTGTCTAAGCCAGAACCGCTCAGTCTGGGCCCTTCTGTGATGCT 338
215 CTGAGAGCTGAGCGGCCAACTGGAGACATATCTGAGCACATATCTGTGGACAATAACC 274
339 CTGAGAGCTGAGCGGCCAACTGGAGACATATCTGAGCACATATCTGTGGACAATAACC 398
275 AGGGGGCCCCGGCAGAGTGGGCACAGGGTGAGCCGCTGAACCGAGAGTCAGAGA 334
399 AGGGGGCCCCGGCAGAGTGGGCACAGGGTGAGCCGCTGAACCGAGAGTCAGAGA 458
335 AGTCCCGGACCTATGTGGCAAGGAATGGGGAGCCCTGAACCAACTCCAGTAGTCAATGGAG 394
459 AGTCCCGGACCTATGTGGCAAGGAATGGGGAGCCCTGAACCAACTCCAGTAGTCAATGGAG 518
395 AGAAGAACCCCTCAGAGGGGATCCAAACACAGAGAGATCCGGCAGAGTAGACAGGTGC 454
519 AGAAGAACCCCTCAGAGGGGATCCAAACACAGAGAGATCCGGCAGAGTAGACAGGTGC 578
455 GAGACGAGACCATCGAAGGCCACAGGAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 514
579 GAGACGAGACCATCGAAGGCCACAGGAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 638
515 TCAGTTGTGTGATGAGACATTTGAATCTCTGAGTACCCGAGAGAGAGTGGGTGCTC 574
639 TCAGTTGTGTGATGAGACATTTGAATCTCTGAGTACCCGAGAGAGAGTGGGTGCTC 698
575 TGTGCAAGAGTATCTGTAAGTCTGGAGGAGCAGCGGAATTCACAGAGCAGATGAAGC 634
699 TGTGCAAGAGTATCTGTAAGTCTGGAGGAGCAGCGGAATTCACAGAGCAGATGAAGC 758
635 TCCTACAGAAAAGCAGAGCCAGTGGTGCAAGAGAGAGCACCACCTGCGCGGTGAGCACA 694
759 TCCTACAGAAAAGCAGAGCCAGTGGTGCAAGAGAGAGCACCACCTGCGCGGTGAGCACA 818
695 GCAAGCGGTCTCTGCGCGCAGCAGCTTGAGAGCCTATGCGGTAGCTGACGCGGACACA 754
819 GCAAGCGGTCTCTGCGCGCAGCAGCTTGAGAGCCTATGCGGTAGCTGACGCGGACACA 878
755 ACCGTCTCTCAGGAAGAGTGTGACGCGGGCCGGAGGAGGAGAGAGGAGGAGGAGGAGG 814
879 ACCGTCTCTCAGGAAGAGTGTGACGCGGGCCGGAGGAGGAGGAGGAGGAGGAGGAGG 938
815 AGTGACCTCGCATTCCAGGTGACACTGAATGACATTCAGTGCAGATGGAACAGCACA 874


```
Db 61 AGCGCGCTCTGTCAGTAGAAGCAGAGAGTCCCGGACAGCAGCAGGCTCTCCGGAAGCCGG 120
QY 167 AGGGTCTCAAGCAGAAACGGCTCAGTCTGGGGCCCTTCGTGATGCTGTCTGAGGAGCTGA 226
Db 121 AGGGGCTCAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTGTCTGAGGAGCTGA 180
QY 227 GCGCCAACTGGAAGACATATCTAGACACATATCTGTGTGACAAATAACCAAGGGGGCCCGG 286
Db 181 GCGCCAACTGGAAGACATATCTAGACACATATCTGTGTGACAAATAACCAAGGGGGCCCGG 240
QY 287 GCGAGATGGGCACAGGCTGAGCGGCTGAAACCGGAGATGAGAGAGTCCCGGACCT 346
Db 241 GCGAGATGGGCACAGGCTGAGCGGCTGAAACCGGAGATGAGAGAGTCCCGGACCT 300
QY 347 ATGTGCAAGGAATGGGGAGCCTGAAACCAACTCCAGTAGTCAATGGAGAGAGGAACCT 406
Db 301 ATGTGCAAGGAATGGGGAGCCTGAAACCAACTCCAGTAGTCTATGGAGAGAGGAACCT 360
QY 407 CCAGGGGATCCAAACAGAGAGATCCGGAGAGTACAGAGTCCGGAGACCGAGACC 466
Db 361 CCAGGGGATCCAAACAGAGAGATCCGGAGAGTACAGAGTCCGGAGACCGAGACC 420
QY 467 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGTTGGGTAAAGAGATCACTTTGCTGA 526
Db 421 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGTTGGGTAAAGAGATCACTTTGCTGA 480
QY 527 TCAGACATTTGAATCTCTGAGTACCCAGAGAGAGAGTGGCTCTCTGTGTCAGAGAGT 586
Db 481 TCAGACATTTGAATCTCTGAGTACCCAGAGAGAGAGTGGCTCTCTGTGTCAGAGAGT 540
QY 587 ATGCTGAACCTGTGGAGGACCGGAATTCACAGAGCAGATGAGCTCTTACAGAAA 646
Db 541 ATGCTGAACCTGTGGAGGACCGGAATTCACAGAGCAGATGAGCTCTTACAGAAA 600
QY 647 AGCAGAGCCAGCTGTGTCAAGAGAGAGCAACCTGCGCGTGCAGCAGCAGAGGCGCTCC 706
Db 601 AGCAGAGCCAGCTGTGTCAAGAGAGAGCAACCTGCGCGTGCAGCAGCAGAGGCGCTCC 660
QY 707 TGGCCGAGCAGAGCTTGAGAGCCTATGCGGTGAGCTGAGCGGCAACACCGCTCCCTCA 766
Db 661 TGGCCGAGCAGAGCTTGAGAGCCTATGCGGTGAGCTGAGCGGCAACACCGCTCCCTCA 720
QY 767 AGAAGAGCTGTGACGCGGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
Db 721 AGAAGAGCTGTGACGCGGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 827 ACTTCAGGTGACACTGAATGACATTCAGTCCAGATGGAACAGCAACATGAGCCCACT 886
Db 781 ACTTCAGGTGACACTGAATGACATTCAGTCCAGATGGAACAGCAACATGAGCCCACT 840
QY 887 CCAAGCTGCGCAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGT 946
Db 841 CCAAGCTGCGCAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
QY 947 ATGAGTGGCGGAGGAGCATATCGAAGAGTCTTCAAAACAGAGGACCTTACAAAGCAGC 1006
Db 901 ATGAGTGGCGGAGGAGCATATCGAAGAGTCTTCAAAACAGAGGACCTTACAAAGCAGC 960
QY 1007 TGGTGGATCCAGCTCCAGAGGCGCGGAGAGATGCTAAAGAGGAGAGAGAGGCGGAC 1066
Db 961 TGGTGGATCCAGCTCCAGAGGCGCGGAGAGATGCTAAAGAGGAGAGAGGCGGAC 1020
QY 1067 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGTATGAGTCCAGAGGATGTGTGAGCTGA 1126
Db 1021 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGTATGAGTCCAGAGGATGTGTGAGCTGA 1080
QY 1127 TGAAGCAGCAGAGAGACCACTGAAAGCAGACAGCTTGCCTTATACACAGAGAGTTGAGG 1186
Db 1081 TGAAGCAGCAGAGAGACCACTGAAAGCAGACAGCTTGCCTTATACACAGAGAGTTGAGG 1140
QY 1187 AGTTCCAGACACATTTCCAAAGCAGCGAGTATTCACCACTTCAAGCAGGAGATGG 1246
```

```
Db 1141 AGTTCCAGAACACACTTTTCCAAAAGCAGCGAGGTATTCCACCATTCAAGCAGGAGATGG 1200
QY 1247 AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGGT 1306
Db 1201 AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGGT 1260
QY 1307 GGGAGAGCAGCAACAAAGCCCTGCTTGAATGGCTGAGGAGAAACAGTCCGGGATAAAG 1366
Db 1261 GGGAGAGCAGCAACAAAGCCCTGCTTGAATGGCTGAGGAGAAACAGTCCGGGATAAAG 1320
QY 1367 AACTGGAGGGCTCAGGTAAATCCAAAGCTGGAGAGAGCTGCGGGCAGTGCAGAG 1426
Db 1321 AACTGGAGGGCTCAGGTAAATCCAAAGCTGGAGAGAGCTGCGGGCAGTGCAGAG 1380
QY 1427 CAGAGCCCAATGACCTGAACAAAGAGGTACAGGACCTGAGTGTGTGGCCAGGCGCTCCC 1486
Db 1381 CAGAGCCCAATGACCTGAACAAAGAGGTACAGGACCTGAGTGTGTGGCCAGGCGCTCCC 1440
QY 1487 TCATGACAGTGGCCCTGAGAGGAGGCGAGAGGCGCTGGGGCTCAAGACCCAGCTCCC 1546
Db 1441 TCRCTGACAGTGGCCCTGAGAGGAGGCGAGAGGCGCTGGGGCTCAAGACCCAGCTCCC 1500
QY 1547 CCAGGCTCACAGAGAGCGCTTGTACCCAGAGCAGCAGCAGCAGAGAGCATCAGGCCAGA 1606
Db 1501 CCAGGCTCACAGAGAGCGCTTGTACCCAGAGCAGCAGCAGCAGCAGAGAGCATCAGGCCAGA 1560
QY 1607 CTGGGCTTCAAGAGCCCACTTCCGCCAGGGCC 1638
Db 1561 CTGGGCTTCAAGAGCCCACTTCCGCCAGGGCC 1592

RESULT 8
AR409323
LOCUS AR409323 4697 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 17 from patent US 6632923.
ACCESSION AR409323
VERSION AR409323.1 GI:40160111
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4697)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 17 14-OCT-2003;
FEATURES
source Location/Qualifiers
1..4697
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 87.9%; Score 1439; DB 6; Length 4697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 47 CAAAAGCAGCCAGGCAACCGGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCC 106
Db 1 CAAAAGCAGCCAGGCAACCGGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCC 60
QY 107 AGCGGCTCTGAGTAGAAGAGAGAGTCCCGGACAGCAGCAGGCTCTTCGGAAGCCGG 166
Db 61 AGCGGCTCTGAGTAGAAGAGAGAGTCCCGGACAGCAGGCTCTTCGGAAGCCGG 120
QY 167 AGGTGCTCAAGCAGAACCGCTCAGTCTGGGCGCTTCGTGATGTCTCTGAGAGCTGA 226
Db 121 AGGGGCTCAAGCAGAACCGCTCAGTCTGGGCGCTTCGTGATGTCTCTGAGAGCTGA 180
QY 227 GCGCCCAACTGGAGACATATCTGAGCACAATCTGTGTGGAACAATAACAGGGGGGCCCG 286
Db 181 GCGCCCAACTGGAGACATATCTGAGCACAATCTGTGTGGAACAATAACAGGGGGGCCCG 240
QY 287 GCGAGGATGGGCAAGGCTGAGCGGCTGAAACCGGAGATGAGAGAGTCCCGGACCT 346
```


Db	241	GCAGAGATGGGCA	CAGGCTGAGCGGCTGAACCCGAAGTGCAGAGAAGTCCCGACCT	300
Qy	347	ATTTGGCAAGAAAT	GGGGAGCCTTGAACCAATCTCAGTAGTCAATGGAGAGAGAAGCAACCT	406
Db	301	ATTTGGCAAGAAAT	GGGGAGCCTTGAACCAATCTCAGTAGTCTATGGAGAGAAGGAACCT	360
Qy	407	CCAAGGGGATCCAA	CACAGAAAGATCCGGCAGAGTGCAGAGTTCGAGAGCCGAGACC	466
Db	361	CCAAGGGGATCCAA	CACAGAAAGATCCGGCAGAGTGCAGAGTTCGAGAGCCGAGACC	420
Qy	467	ATCGAAGGCCACA	GAGGAGAAAAAAGCCAAAGGTTTGGGTAAAGAGATCAAGTTGCTGA	526
Db	421	ATCGAAGGCCACA	GAGGAGAAAAAAGCCAAAGGTTTGGGTAAAGAGATCAAGTTGCTGA	480
Qy	527	TGCAGACATTGAAT	ACTCTCAGTAGTACCCACAGAGAGAGCTGGCTGCTCTGTGCAAGAACT	586
Db	481	TGCAGACATTGAAT	ACTCTCAGTAGTACCCACAGAGAGAGCTGGCTGCTCTGTGCAAGAACT	540
Qy	587	ATGCTGAATGCTG	GAGGAGCACCGAAATTCACAGAAGCAGATGAAGCTCTTACAGAAA	646
Db	541	ATGCTGAATGCTG	GAGGAGCACCGAAATTCACAGAAGCAGATGAAGCTCTTACAGAAA	600
Qy	647	AGCAGGCCAGCTG	GTGTCANAGAGGACACACTGCGCGGTGAGCAGACAGAGCCGCTCC	706
Db	601	AGCAGGCCAGCTG	GTGTCANAGAGGACACACTGCGCGGTGAGCAGACAGAGCCGCTCC	660
Qy	707	TGGCCCGCAGCA	AGCTTTGAGAGCCTATGCGGTGAGCTTGACGCGGCACAACCGTCCCTCA	766
Db	661	TGGCCCGCAGCA	AGCTTTGAGAGCCTATGCGGTGAGCTTGACGCGGCACAACCGTCCCTCA	720
Qy	767	AGGAAGAAGTGTC	AGCGGCGCGGAGAGAGAGAGAGCGCAAGCGCAGAGAGTGACCTCG	826
Db	721	AGGAAGAAGTGTC	AGCGGCGCGGAGAGAGAGAGAGCGCAAGCGCAGAGAGTGACCTCG	780
Qy	827	ACTTTCAGGTGAC	ACTGAATTCAGCTGCAGATGGAACACAGCAATGAGCGCAACT	886
Db	781	ACTTTCAGGTGAC	ACTGAATTCAGCTGCAGATGGAACACAGCAATGAGCGCAACT	840
Qy	887	CCAAGCTGCGCC	CAAGAGAAATGGAGCTGCTGAGAGGCTCAAGAAGCTGATTGAGCACT	946
Db	841	CCAAGCTGCGCC	CAAGAGAAATGGAGCTGCTGAGAGGCTCAAGAAGCTGATTGAGCACT	900
Qy	947	ATGAGCTGCGCG	AGGAGCATATCGACAAAGCTTTCAAACACAAAGGACCTACAACAGCAGC	1006
Db	901	ATGAGCTGCGCG	AGGAGCATATCGACAAAGCTTTCAAACACAAAGGACCTACAACAGCAGC	960
Qy	1007	TGTTGGATGCCA	AGCTCCAGCGCCACAGAGATGCTTAAAGAGGCGAGAGAGCGGCACC	1066
Db	961	TGTTGGATGCCA	AGCTCCAGCGCCACAGAGATGCTTAAAGAGGCGAGAGAGCGGCACC	1020
Qy	1067	AGCGGAGAGAGT	TTTCTCTGAAAGAGGACGTAGAGTCCACAGAGATGTGTGAGCTGA	1126
Db	1021	AGCGGAGAGAGT	TTTCTCTGAAAGAGGACGTAGAGTCCACAGAGATGTGTGAGCTGA	1080
Qy	1127	TGAAGCAGCAGAG	ACCCACTGAAGCAACAGCTTGCCCTTATACAAGAGAACTTTGAGG	1186
Db	1081	TGAAGCAGCAGAG	ACCCACTGAAGCAACAGCTTGCCCTTATACAAGAGAACTTTGAGG	1140
Qy	1187	AGTTCCAGAACAC	ACTTTCCAAAGCAGCAGGTATTCACCACTTCAAGCAGAGAGATGG	1246
Db	1141	AGTTCCAGAACAC	ACTTTCCAAAGCAGCAGGTATTCACCACTTCAAGCAGAGAGATGG	1200
Qy	1247	AAAAGATGACTTA	AGAAGATCAAGAAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGT	1306
Db	1201	AAAAGATGACTTA	AGAAGATCAAGAAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGT	1260
Qy	1307	GGGAGACGCAAC	AAAGCCCTGTTTCAAGTGGCTGAGGAGAAACAGTCCCGGATAAG	1366
Db	1261	GGGAGACGCAAC	AAAGCCCTGTTTCAAGTGGCTGAGGAGAAACAGTCCCGGATAAG	1320
Qy	1367	AAC	TGGAGGGCCTTCAGAGTAAAAATCAACGGCTGGAGAAAGCTGTGCCGGGCACTGCAGA	1426

Db	1321	AACTGAGGGCCCTGCAGGTAAATAATCAACCGGCTGAGGAAGCTGTGCCGGGCACTGCAGA	1388
Qy	1427	CAGAGCGCAATGACCTGAACAACAGAGGGTACAGGACCTGAGTGTGCTGTGGCCAGGGCTCCC	1486
Db	1381	CAGAGCGCAATGACCTGAACAACAGAGGGTACAGGACCTGAGTGTGCTGTGGCCAGGGCTCCC	1440
Qy	1487	TCACTGACAGTGCCCTTGAGAGGAGGCCAGAGGGGCTTGGGGCTCAAGCHACCCAGTCCC	1546
Db	1441	TCACTGACAGTGCCCTTGAGAGGAGGCCAGAGGGGCTTGGGGCTCAAGCAACCAGTCCC	1500
Qy	1547	CCAGGGTCAAGAAAGCGCTTGCTACCCAGGAGCACCCGAGCACAGAAGCATCAGGCCAGA	1606
Db	1501	CCAGGGTCAAGAAAGCGCTTGCTACCCAGGAGCACCCGAGCACAGAAGCATCAGGCCAGA	1560
Qy	1607	CTGGGGCTCAAGAGGCCCACTCCGCCAGGGCC	1638
Db	1561	CTGGGGCTCAAGAGGCCCACTCCGCCAGGGCC	1592
RESULT 9			
AX239574			
LOCUS			
DEFINITION			
AX239574			
ACCESSION			
VERSION			
AX239574.1			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Lees A.M., Lees R.S., Law S.W. and Arjona A.A.			
Low density liprotein binding proteins and their use in diagnosing			
and treating atherosclerosis			
JOURNAL			
Patent: WO 0164874-A 17 07-Sep-2001;			
Boston Heart Foundation, Inc. (US)			
FEATURES			
source			
1. .4697			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
<3..1595			
/notes="unnamed protein product"			
/codon_start=1			
/protein_id="CAC88297.1"			
/db_xref="GI:15797260"			
/db_xref="REMYEMBL:CAC8297"			
/translation="KSPQPEAGPSGAQERPSQAAPAVEAGPGSSQAPRKPEGAQA			
RTAQSGALRDVSELRSQEDILITCYDNNQGGPDGAGQEPAPDAEAKSRVTVA			
RNGEPTPVYGEKPSGKDNTPEIRISQVDSNGDGRDHPQPKKAKGLGKELITLM			
QTANTLSPEELALCKHYALELHSHNSQMKLLQKQSQLVKQKHDLRGHSHKA			
VLARSLGSLCKRLQNSLKEEGVQRAEERKEKVTSHFVTINDILQLOMEHLK			
ERNKLQRNDEMLAEHLKKLIEQVELREEHIDRVFKHDLQQLVDAKLQQAQMLKE			
ABERHQRDFLLKAEVSRMCELMKQETHLQQLALYTEKFEFFQNTLSKSSVFF			
TTPSQWEMXWTKIKLEKETTWTYRSWESSKALLEMAEKTVRKELEGLQVKIQ			
LTKFCRALQTERNDLNKRVQDLSAGGQSLTDSGPERPGPQAQAPSSRVTEAPCY			
FGAPLSTASGQTQGFQPTSARA"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1589; Conservative			
87.9%; Score 1439; DB 6; Length 4697;			
99.8%; Pred. No. 0;			
0; Mismatches 3; Indels 0; Gaps 0;			
Qy	47	CAAAAAGCAGCCAGGACACCGGAGCAGGACCCGAGGAGCCAGGAGCGCCAGGCC	106
Db	1	CAAAAAGCAGCCAGGACACCGGAGCAGGACCCGAGGAGCCAGGAGCGCCAGGCC	60
Qy	107	AGGCGGCTCTCGAGTAGAAGAGCAGAGGTCCTCCGGCAGCAGCCAGGTCCTCTCGAAGCGG	166
Db	61	AGGCGGCTCTCGAGTAGAAGAGCAGAGGTCCTCCGGCAGCAGCCAGGTCCTCTCGAAGCGG	120
Qy	167	AGGCTGTCGAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGA	226

Db 121 AGGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTGAGGAGTGA 180
QY 227 GCGCCCACTGAAGACATACCTAGCAGACATCTGTGTGAGACATTAACAGGGGGGCCCG 286
Db 181 GCGGCCAAGTGAAGACATACCTAGCAGACATCTGTGTGAGACATTAACAGGGGGGCCCG 240
QY 287 GCGAGATGGGGCAGAGGGTGAAGGGTGAACCCGAGATGACAGAGAGTCCCGGAGCT 346
Db 241 GCGAGATGGGGCAGAGGGTGAAGGGTGAACCCGAGATGACAGAGAGTCCCGGAGCT 300
QY 347 ATGTGCAAGGAATGGGGAGCTGAAACCACTCCAGTAGTCAATGAGAGAGAGAACCT 406
Db 301 ATGTGCAAGGAATGGGGAGCTGAAACCACTCCAGTAGTCAATGAGAGAGAGAACCT 360
QY 407 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGAGTGAAGAGTCCGAGACCGAGACC 466
Db 361 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGAGTGAAGAGTCCGAGACCGAGACC 420
QY 467 ATCGAAGGCCACAGGAG 526
Db 421 ATCGAAGGCCACAGGAG 480
QY 527 TCGAGACATGATCTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586
Db 481 TCGAGACATGATCTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 587 ATGCTGAATGCTGGAGGAGACCCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 646
Db 541 ATGCTGAATGCTGGAGGAGACCCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 600
QY 647 AGCAGAGCAGCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
Db 601 AGCAGAGCAGCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 707 TGGCCCGCAGCAAGCTTGAGAGCCTATGCGGTGAGCTGAGCGGACCAACCGCTCCCTCA 766
Db 661 TGGCCCGCAGCAAGCTTGAGAGCCTATGCGGTGAGCTGAGCGGACCAACCGCTCCCTCA 720
QY 767 AGAAGAGAGTGTGAGCGGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
Db 721 AGAAGAGAGTGTGAGCGGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 827 ACTTCAGGTGACATGATGATCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 886
Db 781 ACTTCAGGTGACATGATGATCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 840
QY 887 CCAAGTGGCCCAAGAGAACATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 946
Db 841 CCAAGTGGCCCAAGAGAACATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 900
QY 947 ATGAGTGGCGAGAGAGATATCGAAGAGTCTTCAACAGAGAGAGAGAGAGAGAGAG 1006
Db 901 ATGAGTGGCGAGAGAGATATCGAAGAGTCTTCAACAGAGAGAGAGAGAGAGAGAG 960
QY 1007 TGGTGTATCCAGTCCAGAGAGCCAGAGATGCTAAGAGAGAGAGAGAGAGAGAGAG 1066
Db 961 TGGTGTATCCAGTCCAGAGAGCCAGAGATGCTAAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1067 AGCGGAGAGAGATTTCTTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
Db 1021 AGCGGAGAGAGATTTCTTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1127 TGAAGCAGACAG 1186
Db 1081 TGAAGCAGACAG 1140
QY 1187 AGTTCAGAGACACTTTTCTTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
Db 1141 AGTTCAGAGACACTTTTCTTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1247 AAAAGATGATGAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
Db 1201 AAAAGATGATGAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

QY 1307 GGGAGACAGCAACAGGCCCTCTTGTAGATGCTGAGGAGAGAGAGAGAGAGAGAG 1366
Db 1361 GGGAGACAGCAACAGGCCCTCTTGTAGATGCTGAGGAGAGAGAGAGAGAGAGAG 1320
QY 1367 AACTGGAGGGCCCTGCGAGTAAATCAACGGCTGGAGAGAGAGAGAGAGAGAG 1426
Db 1321 AACTGGAGGGCCCTGCGAGTAAATCAACGGCTGGAGAGAGAGAGAGAGAGAG 1380
QY 1427 CAGAGCCATGACCTGAACAGAGGGTACAGAGACCTGAGTGTGTGGCCAGGGCTCCC 1486
Db 1381 CAGAGCCATGACCTGAACAGAGGGTACAGAGACCTGAGTGTGTGGCCAGGGCTCCC 1440
QY 1487 TCACTGACAGTGGCCCTGAGAGAGGGCCAGAGGGCTGAGGGCTCAAGCACCCAGCTCCC 1546
Db 1441 TCACTGACAGTGGCCCTGAGAGAGGGCCAGAGGGCTGAGGGCTCAAGCACCCAGCTCCC 1500
QY 1547 CCAGGGTCAAGAGAGCCCTTGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1606
Db 1501 CCAGGGTCAAGAGAGCCCTTGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1607 CTGGGGCTCAAGAGAGCCACCTCCGCCAGGGGCC 1638
Db 1561 CTGGGGCTCAAGAGAGCCACCTCCGCCAGGGGCC 1592

RESULT 10
LOCUS BD056450 4697 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis,
ACCESSION BD056450
VERSION BD056450.1 GI:22602056
KEYWORDS JP 2001506983-A/8.
SOURCE Aequorea victoria
ORGANISM Aequorea victoria
Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Leptomedusae;
Aequoreidae; Aequorea.
REFERENCE 1 (bases 1 to 4697)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Novel low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: JP 2001506983-A 8 29-MAY-2001;
BOSTON HEART FOUNDATION INC
COMMENT PN JP 2001506983-A/8
PD 29-MAY-2001
PF 26-NOV-1997 JP 1998524870
PR 27-NOV-1996 US 60/031930, 03-JUN-1997 US 60/048547 PI
ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC
A61K38/04, A61K38/17, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC
C07H21/00,
PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
source
1..4697
/organism="Aequorea victoria"
/mol_type="Genomic DNA"
/db_xref="taxon:6100"

Query Match 87.9%; Score 1439; DB 6; Length 4697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 47 CAAAAGCAGCCAGGACAAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 106
Db 1 CAAAAGCAGCCAGGACAAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 107 AGCGGGCTCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166
Db 61 AGCGGGCTCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

```
QY 167 AGGGTGTCTCAAGCCAGAACGGCTCAGTCTGTGGGGCCCTTCTGTGATGTCTCTGAGGAGCTGA 226
DB 121 AGGGGGCTCAAGCCAGAACGGCTCAGTCTGTGGGGCCCTTCTGTGATGTCTCTGAGGAGCTGA 180
QY 227 GCGCCAACTGGAAGACATACCTAGAGACATATCTGTGTGGACAAATTAACCGGGGGGCCCCG 286
DB 181 GCGCCAACTGGAAGACATACCTAGAGACATATCTGTGTGGACAAATTAACCGGGGGGCCCCG 240
QY 287 GCGAGGATCGGGCACAGGGTGAAGCGGCTGAACCGGAGATCAGAGAACTCCCGGACT 346
DB 241 GCGAGGATCGGGCACAGGGTGAAGCGGCTGAACCGGAGATCAGAGAACTCCCGGACT 300
QY 347 ATGTGCAAGGAATCGGGAGCTTGAACCAACTCCAGTAGTCAATGGAGAGAAAGCAACCT 406
DB 301 ATGTGCAAGGAATCGGGAGCTTGAACCAACTCCAGTAGTCAATGGAGAGAAAGCAACCT 360
QY 407 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTCAACAGGTCCGAGACCCGAGACC 466
DB 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTCAACAGGTCCGAGACCCGAGACC 420
QY 467 ATCGAAGGCCACAGGAGAGAGAAAAAAGCCAAAGGGTTTGGTAAAGAGATCACGTTGCTGA 526
DB 421 ATCGAAGGCCACAGGAGAGAGAAAAAAGCCAAAGGGTTTGGGGAAGAGATCACGTTGCTGA 480
QY 527 TGCAAGATTTGAATATCTCTGAGTACCCAGAGAGAGAGCTGCTGCTGTGCAAGAGT 586
DB 481 TGCAAGATTTGAATATCTCTGAGTACCCAGAGAGAGAGCTGCTGCTGTGCAAGAGT 540
QY 587 ATGCTCAACTGCTGGAGGAGACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAAA 646
DB 541 ATGCTCAACTGCTGGAGGAGACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAAA 600
QY 647 AGCAGAGCAGTGTGTGCAAGAGAGAGAACCACTGCGCGGTGAGACACAGCAAGCGCGTCC 706
DB 601 AGCAGAGCAGTGTGTGCAAGAGAGAGAACCACTGCGCGGTGAGACACAGCAAGCGCGTCC 660
QY 707 TGCGCCGCGAGCAGCTTGAAGCCTATGCGGTGAGCTGCGAGCGGACACACCGCTCCCTCA 766
DB 661 TGCGCCGCGAGCAGCTTGAAGCCTATGCGGTGAGCTGCGAGCGGACACACCGCTCCCTCA 720
QY 767 AGGAAGAAGTGTGTGACGCGGCGCGGAGAGAGAGAACCGCAAGAGGTGACCTGCG 826
DB 721 AGGAAGAAGTGTGTGACGCGGCGCGGAGAGAGAGAACCGCAAGAGGTGACCTGCG 780
QY 827 ACTTCAGGTGACATCACTGATGACATTCAGCTGAGTGAAGCAGACATGAGCGCAACT 886
DB 781 ACTTCAGGTGACATCACTGATGACATTCAGCTGAGTGAAGCAGACATGAGCGCAACT 840
QY 887 CCAAGCTGCGCCAAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 946
DB 841 CCAAGCTGCGCCAAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 900
QY 947 ATGAGCTGCGCGAGGAGCATATGCAAAAGTCTTCAACACAGAGACCTACAGAGCAGC 1006
DB 901 ATGAGCTGCGCGAGGAGCATATGCAAAAGTCTTCAACACAGAGACCTACAGAGCAGC 960
QY 1007 TGTGTGATGCAAGCTCCAGCAGGCGCCAGAGATGCTAAAGAGGCGAGAGCGCGCACC 1066
DB 961 TGTGTGATGCAAGCTCCAGCAGGCGCCAGAGATGCTAAAGAGGCGAGAGCGCGCACC 1020
QY 1067 AGCGGAGAGAGATTTCTTCTGAAAGAGCGAGTAGTCCAGAGGATGTGTGAGCTGA 1126
DB 1021 AGCGGAGAGAGATTTCTTCTGAAAGAGCGAGTAGTCCAGAGGATGTGTGAGCTGA 1080
QY 1127 TGAAGCAGAGAGAGACCCACCTGGAACACAGCTTGCCCTATACACAGAGAGTTTGAGG 1186
DB 1081 TGAAGCAGAGAGAGACCCACCTGGAACACAGCTTGCCCTATACACAGAGAGTTTGAGG 1140
QY 1187 AGTTCCAGAACACATTTTCCAAAGCAGCGAGTATTCACCACTTCAAGCAGGAGATGG 1246
DB 1141 AGTTCCAGAACACATTTTCCAAAGCAGCGAGTATTCACCACTTCAAGCAGGAGATGG 1200
```

```
QY 1247 AAAAGATGACTAAGAGATCAAGAGCTGAGAGAAAGAAACCACTGTACCGGTCCCGGT 1306
DB 1201 AAAAGATGACTAAGAGATCAAGAGCTGAGAGAAAGAAACCACTGTACCGGTCCCGGT 1260
QY 1307 GGGAGAGCAGCAACAGGCCCTCTGTGATGCTGAGGAGAAACAGTCCGGGATTAAG 1366
DB 1261 GGGAGAGCAGCAACAGGCCCTCTGTGATGCTGAGGAGAAACAGTCCGGGATTAAG 1320
QY 1367 AACTGGAGGGCTCGCAGGTAAAAATCCACGGCTGGAGAGCTGTGCCGGGCACTCCAGA 1426
DB 1321 AACTGGAGGGCTCGCAGGTAAAAATCCACGGCTGGAGAGCTGTGCCGGGCACTCCAGA 1380
QY 1427 CAGAGCCCATGACTGAACAGAGGGTACAGAGCTGAGTCTGTGGCCAGCGGTCC 1486
DB 1381 CAGAGCCCATGACTGAACAGAGGGTACAGAGCTGAGTCTGTGGCCAGCGGTCC 1440
QY 1487 TCACTGACGTGGCCCTGAGAGGAGCCAGAGGGGCTCGGGCTCAAGCACCCAGCTCC 1546
DB 1441 TCACTGACGTGGCCCTGAGAGGAGCCAGAGGGGCTCGGGCTCAAGCACCCAGCTCC 1500
QY 1547 CAGGGTCAAGAGCGCTTGTACCCAGGAGCACCAGGACACAGAGCATCAGGCCAGA 1606
DB 1501 CAGGGTCAAGAGCGCTTGTACCCAGGAGCACCAGGACACAGAGCATCAGGCCAGA 1560
QY 1607 CTGGGCTCAAGAGCCACCTCCCGCCAGGGCC 1638
DB 1561 CTGGGCTCAAGAGCCACCTCCCGCCAGGGCC 1592

RESULT 11
HSM803947 5023 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKF2p45J0118 (from clone DKF2p45J0118).
DEFINITION AL832637
ACCESSION AL832637.2 GI:30268326
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 5023)
AUTHORS Ansoorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg GERMANY
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733213.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKF2p45J0118) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..5023
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKF2p45J0118"
/db_xref="taxon:9606"
/map="lp34.2-36.11"
/clone="DKF2p45J0118"
/tissue_type="human spinal cord"
/clone_lib="451 (synonym: hlcc1). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..5023
/gene="DKF2p45J0118"
/join(283..696,696..1922)

gene
CDS
```

```

/gene="DKF2p451J0118"
/Note="hypothetical protein
The frame shift was determined manually"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD89952.1"
/db_xref="GI:30268327"
/translation="MKNDKNGAKQSNPKSPQEPBAGPAGPQPSQAPAVEAE
GPGSQAARKPEGAQTAAQSGALRDVSEELSRQLEDILSTYCVNNQGGPGEAGQ
EPBPEDAEKRTYVARGNEPPTPVNGEKEPKGDPNTEEIFRQSEVGGDRDHRPO
BKAKAGTGGTLLMOTLNTLSPEEKLAALCRKYABLLSEHNSOKMKLLKKOS
QLVQKDLHREHSAVLARSLSLRELOHNRSLKEEGVQBARSEEEKKRVTH
FLVTLNDLQLOMEQNERNSKLQRNEMLAERLKLIEQYELREHIDKVPKDLQO
QLVDKLAQQAQMEKEAEERHOREKDFLLKEAVASORMCELMKQETHLKKQALYTE
KFEPQNTLSKSEVFTFKQEMKVTTKIKLEKETTYMRSESNKALLEAEK
TVRDELGLQVIOIRLEKLCRALOTERNLNRVODLSAGGQGLTDSGPERPEGP
GAQAPSSPRVTEAPCPGAPSTEASGQTGPQEPITSARA"
4991..4996
polyA_signal
/gene="DKF2p451J0118"
5010
polyA_site
/gene="DKF2p451J0118"
ORIGIN

```

```

Query Match      77.0%; Score 1262; DB 9; Length 5023;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1 ATGAAGAACCAAGACAAAGAAAGACGGGGCTGCCAAACCAATCCAAATCCAAAAGACGACCCCA 60
DB 283 ATGAAGAACCAAGACAAAGAAAGACGGGGCTGCCAAACCAATCCAAATCCAAAAGACGACCCCA 342
QY 61 GGACAAACCGAGACGAGACCCGAGGAGCCGAGGAGCGGCGCCAGCAGCGCGCTCTTGA 120
DB 343 GGACAAACCGAGACGAGACCCGAGGAGCCGAGGAGCGGCGCCAGCAGCGCGCTCTTGA 402
QY 121 GTAGAAGACGAGAGTCCCGGACGACGACGAGGCTCTCGGAGCCGAGGAGGCTCTCAAGCC 180
DB 403 GTAGAAGACGAGAGTCCCGGACGACGACGAGGCTCTCGGAGCCGAGGAGGCTCTCAAGCC 462
QY 181 AGAAAGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGAGCGCGCAACTGAA 240
DB 463 AGAAGCGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGAGCGCGCAACTGAA 522
QY 241 GACATAGTACGACATCTGTGTGACAAATACCAAGGGGGGCGCGGAGGATGGGGCA 300
DB 523 GACATAGTACGACATCTGTGTGACAAATACCAAGGGGGGCGCGGAGGATGGGGCA 582
QY 301 CAGGCTGAGCCGGCTGAACCCGAGATGCGAGAAAGTCCCGGACCTATGTGCAAGGAAT 360
DB 583 CAGGCTGAGCCGGCTGAACCCGAGATGCGAGAAAGTCCCGGACCTATGTGCAAGGAAT 642
QY 361 GGGGAGCTTGAAACCACTCAGTAGTCAATGAGAGAGGAGACCTCCCAAGGGGATCCA 420
DB 643 GGGGAGCTTGAAACCACTCAGTAGTCAATGAGAGAGGAGACCTCCCAAGGGGATCCA 701
QY 421 AACACAGAGAGATCCGGCAGAGTGAACGAGGTGCGAGACCGAGACCATCGAAGGCCACAG 480
DB 702 AACACAGAGAGATCCGGCAGAGTGAACGAGGTGCGAGACCGAGACCATCGAAGGCCACAG 761
QY 481 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 762 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
QY 541 ACTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 822 ACTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
QY 601 GAGGAGACCCGGAATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 882 GAGGAGACCCGGAATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 941
QY 661 GTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720

```

```

RESULT 12
AX578043/c
LOCUS
DEFINITION Sequence 165 from Patent WO02081745.
ACCESSION AX578043
VERSION AX578043.1 GI:27647251

```

```

AX578043 1793 bp DNA linear PAT 08-JAN-2003

```


||||| AGCAATGAGCGCAATCCAAAGCTGCGCCAAAGAAATGAGCTGGCTGAGAGGCTCA 1477
QY 929 AGAGCTGATTGACAGTATGAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACA 988
DB 1476 AGAAGCTGATTGACAGTATGAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACA 1417
QY 989 AGGACTCAACACAGCAGCTGGTGGATGTCAGAGCTCCAGCAGCGCCAGAGAGTGTAAAGG 1048
DB 1416 AGGACCTACACAGCAGCTGGTGGATGTCAGAGCTCCAGCAGCGCCAGAGAGTGTAAAGG 1357
QY 1049 AGGAGAGAGGGGACAGCGGGAGAGAGATTTCTCTCTGAAGAGGAGTGTAGTCCC 1108
DB 1356 AGGCAAGAGCGGACAGCGGGAGAGAGATTTCTCTCTGAAGAGGAGTGTAGTCCC 1257
QY 1109 AGAGATGCTGTGAGCTGATGAAGCAGCAGAGAGAGCCCTCAAGCAACAGCTTCCCTAT 1168
DB 1296 AGAGATGCTGTGAGCTGATGAAGCAGCAGAGAGAGCCCTCAAGCAACAGCTTCCCTAT 1237
QY 1169 ACACAGAGAGTTTGAAGAGTTCAGAGACACACTTCCAAAGAGAGGAGTGTATCACCA 1228
DB 1236 ACACAGAGAGTTTGAAGAGTTCAGAGACACACTTCCAAAGAGAGGAGTGTATCACCA 1177
QY 1229 CATTCAGCAGAGATGAAAGATGACTAAGAGATCAAGAGCTGAGAGAGAAACCA 1288
DB 1176 CATTCAGCAGAGATGAAAGATGACTAAGAGATCAAGAGCTGAGAGAGAAACCA 1117
QY 1289 CCATGTACCGGTCCCGGTGGAGAGCAGACACAGCGCCCTCTTGAAGTGTGAGAGAGA 1348
DB 1116 CCATGTACCGGTCCCGGTGGAGAGCAGACACAGCGCCCTCTTGAAGTGTGAGAGAGA 1057
QY 1349 AAACAGTCGGGATAAAGACTGAGAGGCGCTGAGGTAAAATCAACGGCTGGAGAGC 1408
DB 1056 AAACAGTCGGGATAAAGACTGAGAGGCGCTGAGGTAAAATCAACGGCTGGAGAGC 997
QY 1409 TGTGCGGGGCACTGACAGACAGCGCAATGACCTGAACAGAGAGGTACAGACCTGAGTG 1468
DB 996 TGTGCGGGGCACTGACAGACAGCGCAATGACCTGAACAGAGAGGTACAGACCTGAGTG 937
QY 1469 CTGFTGGCCAGGGTCCCTCACTGACAGTGGCCCTTGAGAGAGGCGCAGAGGGGCTGGG 1528
DB 936 CTGFTGGCCAGGGTCCCTCACTGACAGTGGCCCTTGAGAGAGGCGCAGAGGGGCTGGG 877
QY 1529 CTCAGACACAGCTCCCGGCTGACAGAGGCGCTTGCTACCCAGGAGCAGCGAGCA 1588
DB 876 CTCAGACACAGCTCCCGGCTGACAGAGGCGCTTGCTACCCAGGAGCAGCGAGCA 817
QY 1589 CAGAGCATCAGCGCAGAGTGGGCTCAAGAGCGCCACCTCCCGCAGGGGCC 1638
DB 816 CAGAGCATCAGCGCAGAGTGGGCTCAAGAGCGCCACCTCCCGCAGGGGCC 767

RESULT 14
LOCUS BC046565 3554 bp mRNA linear PRI 30-SEP-2003
DEFINITION Homo sapiens taxilin, mRNA (cDNA clone IMAGE:5534975), complete cds
ACCESSION BC046565
VERSION BC046565.1 GI:28302155
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3554)
AUTHORS Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grinwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3554)
Direct Submission
Submitted (10-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: ARCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 78 Row: D Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
This clone has the following problem: The cds is short compared to the longest cds in the locus.
Location/Qualifiers
1. 3554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5534975"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 3554
/genes="DKFZp451J0118"
/db_xref="LocusID:200081"
68_418
/codon_start=1
/product="DKFZp451J0118 protein"
/protein_id="AAH46565.1"
/db_xref="GI:28302156"
/db_xref="LocusID:200081"
/translation="MYRSRWSSNKALLEMAEKTVRDKELEGLOVKIQRLKLCRAL QTERNDLNKSVQDLSAGGGGSLTDSGPFRRPGGAQPSPRVTEAPCVPGAPSTEA SGTGTGQETPSARA"

FEATURES

source

gene

CDS

ORIGIN

Query Match 25.3%; Score 415; DB 9; Length 3554;
 Best Local Similarity 100.0%; Pred. No. 4e-226;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1224 CACCACATTCAAGCAGGAGATGGAAGAAGATGACTAAGAAGATCAAGAAGCTGGAGAAGA 1283
 DB 1 CACCACATTCAAGCAGGAGATGGAAGAAGATGACTAAGAAGATCAAGAAGCTGGAGAAGA 60
 QY 1284 ACCACCATGTACCGTCCCGTGGAGAGCAGCAGCAGGCGCTTGTAGATGGCTGA 1343
 DB 61 ACCACCATGTACCGTCCCGTGGAGAGCAGCAGCAGGCGCTTGTAGATGGCTGA 120
 QY 1344 GGAGAAAACAGTCCCGGATAAAGAACTGGAGGGCCCTGCAGGTAAATAATCCAAACGGCTGA 1403
 DB 121 GGAGAAAACAGTCCCGGATAAAGAACTGGAGGGCCCTGCAGGTAAATAATCCAAACGGCTGA 180
 QY 1404 GAAGCTGTCCGGGCACTGCACAGAGCGCAATGACCTGACAGAGGGTACAGGACCT 1463
 DB 181 GAAGCTGTCCGGGCACTGCACAGAGCGCAATGACCTGACAGAGGGTACAGGACCT 240
 QY 1464 GAGTGTCTGTGCGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGCGCCAGAGGGGCC 1523
 DB 241 GAGTGTCTGTGCGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGCGCCAGAGGGGCC 300
 QY 1524 TGGGGTCAAGCACCAGCTCCCGCAGGGTCAAGAAGCGCTTGTCTACCCAGAGGACCC 1583
 DB 301 TGGGGTCAAGCACCAGCTCCCGCAGGGTCAAGAAGCGCTTGTCTACCCAGAGGACCC 360
 QY 1584 GAGCAGAGGATCAGGCGCAGCTGGGCTCAAGAGCCACCTCCGCGCAGGGCC 1638
 DB 361 GAGCAGAGGATCAGGCGCAGCTGGGCTCAAGAGCCACCTCCGCGCAGGGCC 415

RESULT 15
 BD060509
 LOCUS 529 bp DNA linear PAT 27-AUG-2002
 DEFINITION Secreted expressed sequence tags (sESTs).
 ACCESSION BD060509
 VERSION BD060509.1 GI:22606115
 KEYWORDS JP 2001518793-A/869.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 529)
 Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
 Treacy,M., Spaulding,V. and Agostino,M.J.
 Title Secreted expressed sequence tags (sESTs)
 Patent: JP 2001518793-A 869 16-OCT-2001;
 GENETICS INSTITUTE INC
 FN JP 2001518793-A/869
 PD 16-OCT-2001
 PF 10-APR-1998 JP 1998543070
 PR 10-APR-1997 US 08/837312
 PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
 DAVID MERBERG
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
 C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
 Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 Location/Qualifiers
 1..529
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /db_xref="taxon:4577"

FEATURES
 source
 1..529
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /db_xref="taxon:4577"

ORIGIN
 Query Match 25.2%; Score 413; DB 6; Length 529;
 Best Local Similarity 99.8%; Pred. No. 5.2e-225;
 Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGAACCAAGACAAAAAGAAACGGGGTCTCCAAACAATCCAAATCCAAAAAGCAGCCCA 60
 DB 56 ATGAAGAACCAAGACAAAAAGAAACGGGGTCTCCAAACAATCCAAATCCAAAAAGCAGCCCA 115
 QY 61 GGCACACCGAAGCAGAGACCCGAGGAGGCCAGGAGCGGCCAGGCGGCTCTCTGA 120
 DB 116 GGCACACCGAAGCAGAGACCCGAGGAGGCCAGGAGCGGCCAGGCGGCTCTCTGA 175
 QY 121 GTAGAAGCACAAGGTCCCGCAGCAGCAGGCTCTCGGAAGCGCGAGGCTGTCTCAAGCC 180
 DB 176 GTAGAAGCACAAGGTCCCGCAGCAGCAGGCTCTCGGAAGCGCGAGGCTGTCTCAAGCC 235
 QY 181 AGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTGTAGAGAGCTGAGCGCCCAACTGGAA 240
 DB 236 AGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTGTAGAGAGCTGAGCGCCCAACTGGAA 295
 QY 241 GACATCTGAGCACAATCTGTGTGGACAATAACCCAGGGGGCCCCCGCAGGATGGGCA 300
 DB 296 GACATCTGAGCACAATCTGTGTGGACAATAACCCAGGGGGCCCCCGCAGGATGGGCA 355
 QY 301 CAGGGTCAAGCGGCTGAACCCGAGGATGCGAGAGTCCCGGACCTATGTGCAAGGAAT 360
 DB 356 CAGGGTCAAGCGGCTGAACCCGAGGATGCGAGAGTCCCGGACCTATGTGCAAGGAAT 415
 QY 361 GGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAGAACCTCCCAAGGGGATCCA 420
 DB 416 GGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAGAACCTCCCAAGGGGATCCA 475
 QY 421 AACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGAGCCGAGA 464
 DB 476 AACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGAGCCGAGA 519

RESULT 16
 BC029686
 LOCUS 3495 bp mRNA linear PRI 20-MAY-2002
 DEFINITION Homo sapiens, clone IMAGE:4993857, mRNA, partial cds.
 ACCESSION BC029686
 VERSION BC029686.1 GI:20987361
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3495)
 Strausberg,R.
 Direct Submission
 Submitted (06-MAY-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 42 Row: j Column: 8.
 Location/Qualifiers
 1..3495
 /organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4993857"
/tissue_type="placenta, choriocarcinoma"
/clone_lib="NIH MGC_10"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1.393
/codon_start=1
/product="Unknown (protein for IMAGE:4993857)"
/protein_id="AAH29686.1"
/db_xref="GI:20987362"
/translation="KMTKKIKLEKTTMYSRWSSNKALLEWAEKTVKLEGL
QVKIQRLEKLCRALQTERNDLNKRVQDLGAGCGGSLTDSGPFRRPFGAAPS
TEAPYCPGAPSTASGQTGPBPTSARA"

CDS

ORIGIN

Query Match 23.8%; Score 390; DB 9; Length 3495;
Best Local Similarity 100.0%; Pred. No. 9.5e-212; Indels 0; Gaps 0;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1249 AAGATGACTAAGAGATCAAGAGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 1308
DB 1 AAGATGACTAAGAGATCAAGAGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 60
QY 1309 GAGAGCAGCAACAGCCCTGCTTGATGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 1368
DB 61 GAGAGCAGCAACAGCCCTGCTTGATGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 120
QY 1369 CTGGAGGGGCTGAGGATCAAGAGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 1428
DB 121 CTGGAGGGGCTGAGGATCAAGAGCTGGAGAGAAACACCATGTACCGGTCCCGGTGG 180
QY 1429 GAGCGCAATGACCTGAACAGAGGCTGAGGAGAAACACCATGTACCGGTCCCGGTGG 1488
DB 181 GAGCGCAATGACCTGAACAGAGGCTGAGGAGAAACACCATGTACCGGTCCCGGTGG 240
QY 1489 ACTGACAGTGGGCTGAGGAGGCTGAGGAGAAACACCATGTACCGGTCCCGGTGG 1548
DB 241 ACTGACAGTGGGCTGAGGAGGCTGAGGAGAAACACCATGTACCGGTCCCGGTGG 300
QY 1549 AGGGTCACAGAGAGGCTGCTACCGAGGAGGCTGAGGAGAAACACCATGTACCGGTGG 1608
DB 301 AGGGTCACAGAGAGGCTGCTACCGAGGAGGCTGAGGAGAAACACCATGTACCGGTGG 360
QY 1609 GGGCTCAGAGAGGCTGCTACCGAGGAGGCTGAGGAGAAACACCATGTACCGGTGG 1638
DB 361 GGGCTCAGAGAGGCTGCTACCGAGGAGGCTGAGGAGAAACACCATGTACCGGTGG 390

RESULT 17

AR409342
LOCUS AR409342 22255 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 51 from patent US 6632923.
ACCESSION AR409342
VERSION AR409342.1 GI:40160130
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 22255)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 51 14-OCT-2003;
FEATURES Location/Qualifiers
source
1. .22255
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 20.5%; Score 336; DB 6; Length 22255;
Best Local Similarity 100.0%; Pred. No. 1.2e-180;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 231
DB 2999 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 3058
QY 232 CAACCTGGAAGACATACCTGAGCACAATCTGTGTGCAACAATAACACAGGGGGCCCGGGCGAG 291
DB 3059 CAACCTGGAAGACATACCTGAGCACAATCTGTGTGCAACAATAACACAGGGGGCCCGGGCGAG 3118
QY 292 GATGGGCGACAGGGTGAAGCCGGCTGAACCCGAGAGATGACAGAGAGTCCCGGACCTATGTG 351
DB 3119 GATGGGCGACAGGGTGAAGCCGGCTGAACCCGAGAGATGACAGAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCAATCGAGAGAGAGGACCTCCCAAG 411
DB 3179 GCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCAATCGAGAGAGAGGACCTCCCAAG 3238
QY 412 GGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGACCGAGACCATCGA 471
DB 3239 GGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGACCGAGACCATCGA 3298
QY 472 AGGCCACAG 507
DB 3299 AGGCCACAG 3334
RESULT 18
AR4239608
LOCUS AR4239608 22255 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 51 from Patent WO0164874.
ACCESSION AR4239608
VERSION AR4239608.1 GI:15797283
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 51 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES Location/Qualifiers
source
1. .22255
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 20.5%; Score 336; DB 6; Length 22255;
Best Local Similarity 100.0%; Pred. No. 1.2e-180;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 231
DB 2999 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 3058
QY 232 CAACCTGGAAGACATACCTGAGCACAATCTGTGTGCAACAATAACACAGGGGGCCCGGGCGAG 291
DB 3059 CAACCTGGAAGACATACCTGAGCACAATCTGTGTGCAACAATAACACAGGGGGCCCGGGCGAG 3118
QY 292 GATGGGCGACAGGGTGAAGCCGGCTGAACCCGAGAGATGACAGAGAGTCCCGGACCTATGTG 351
DB 3119 GATGGGCGACAGGGTGAAGCCGGCTGAACCCGAGAGATGACAGAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCAATCGAGAGAGAGGACCTCCCAAG 411
DB 3179 GCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCAATCGAGAGAGAGGACCTCCCAAG 3238
QY 412 GGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGACCGAGACCATCGA 471

Db 3239 GGGGATCCAAACACAGAGATCCGGCAGAGTGACGAGTGGAGACGACCATCGA 3298

QY 472 AGCCACAGAGAGAGAAAAAGCCAAAGGTTTCGGT 507
|||||

Db 3299 AGCCACAGAGAGAGAAAAAGCCAAAGGTTTCGGT 3334
|||||

RESULT 19
HSDJ622L5
LOCUS HSDJ622L5 115756 bp DNA linear PRI 22-NOV-2001
DEFINITION Human DNA sequence from clone RP4-622L5 on chromosome 1p34.2-36.11,
complete sequence.
ACCESSION AL049795
VERSION AL049795.21 GI:17065925
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Frankland, J.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
COMMENT On Nov 25, 2001 this sequence version replaced gi:6010175.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.: EMBL; Sw.: SWISSPROT; Tr.: TrEMBL; Wp.: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone configs of human
Chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
This sequence is the entire insert of clone RP4-622L5. This sequence
was finished as follows unless otherwise noted: all regions were
either double-stranded or sequenced with an alternate chemistry or
covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at least one
plasmid subclone or more than one M13 subclone; and the assembly
was confirmed by restriction digest. RP4-622L5 is from the library
RPCI-4 constructed by the group of Pieter de Jong. For further
details see
http://www.chori.org/bacpac/home.htm
VECTOR: PCYPAC2.

FEATURES
source
1..115756
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="p34.2-36.11"
/clone="RP4-622L5"
/clone_lib="RPCI-4"
1645..1810
/note="TIGGER1 repeat: matches 2259..2417 of consensus"
2123..2437
/note="TIGGER1 repeat: matches 1957..2259 of consensus"
2726..2782
/note="TIGGER1 repeat: matches 1904..1957 of consensus"
3067..3486
/note="TIGGER1 repeat: matches 1501..1904 of consensus"
3781..4180
/note="TIGGER1 repeat: matches 1117..1501 of consensus"
5053..5082

repeat_region
/note="TIGGER1 repeat: matches 1090..1117 of consensus"
5372..5501
/note="TIGGER1 repeat: matches 966..1090 of consensus"
5628..5704
/note="TIGGER1 repeat: matches 893..966 of consensus"
5991..6492
/note="TIGGER1 repeat: matches 411..893 of consensus"
6800..6835
/note="TIGGER1 repeat: matches 378..411 of consensus"
6988..7022
/note="TIGGER1 repeat: matches 1..378 of consensus"
8281..8358
/note="3 copies 26 mer 79% conserved"
8301..8360
/note="30 copies 2 mer ta 81% conserved"
17674..17777
/note="4 copies 26 mer 77% conserved"
17885..17784
/note="50 copies 2 mer at 82% conserved"
19385..19542
/note="match: GSS: Em:AQ306911"
19551..19574
/note="12 copies 2 mer tt 100% conserved"
20865..20922
/note="29 copies 2 mer at 70% conserved"
20934..21043
/note="55 copies 2 mer at 70% conserved"
20940..21043
/note="4 copies 26 mer 74% conserved"
21168..21234
/note="tRNA-Asp-GAC repeat: matches 1..75 of consensus"
28263..28691
/note="match: GSS: Em:AQ317905"
complement(23964..30447)
/note="match: GSS: Em:AQ732256"
30465..30718
/note="match: STS: Em:G58909
match: GSS: Em:AQ376964"
38923..39446
/note="L1MEC repeat: matches 1546..2095 of consensus"
39761..39882
/note="L1MEC repeat: matches 1429..1546 of consensus"
complement(43254..43726)
/note="match: GSS: Em:AQ146371"
45555..46050
/note="MER44C repeat: matches 1..728 of consensus"
46055..46468
/note="match: GSS: Em:AQ040362"
48914..48982
/note="MER21B repeat: matches 718..790 of consensus"
48985..49173
/note="MER21B repeat: matches 77..263 of consensus"
49285..49651
/note="match: GSS: Em:AQ761532"
52462..53177
/note="match: GSS: Em:AQ040082"
54716..54865
/note="match: GSS: Em:AQ680771"
60551..60614
/note="32 copies 2 mer tt 70% conserved"
complement(62219..62382)
/note="match: STS: Em:G20585"
complement(66030..66267)
/note="match: GSS: Em:AQ514287"
complement(66041..66267)
/note="match: GSS: Em:AQ803288"
66066..66255
/note="match: GSS: Em:B69239"
68882..69541
/note="CpG island"
71427..71449
/note="11 copies 2 mer tt 100% conserved"


```
misc_feature 73527..74249 /note="match: GSS: Em:AQ892680"
misc_feature 74514..74926 /note="match: GSS: Em:AQ441232"
misc_feature 74532..74981 /note="match: GSS: Em:AQ812740"
misc_feature 74626..75008 /note="match: GSS: Em:AQ091614"
misc_feature complement(74682..75020) /note="match: STS: Em:L29744"
misc_feature complement(74683..74951) /note="match: STS: Em:L29729 Em:L29763"
misc_feature 74717..74867 /note="match: STS: Em:G15744"
misc_feature 82692..83089 /note="match: STS: Em:G09447"
misc_feature 83391..83481 /note="match: GSS: Em:A2310334"
repeat_region 85284..85337 /note="27 copies 2 mer tt 88% conserved"
misc_feature 85339..85718 /note="match: STS: Em:G28385"
misc_feature complement(87250..87612) /note="match: STS: Em:G06968"
misc_feature complement(87455..87579) /note="match: STS: Em:G26511"
misc_feature complement(87831..88313) /note="match: GSS: Em:AQ536603"
repeat_region 89654..89685 /note="16 copies 2 mer tc 87% conserved"
misc_feature 91856..92386 /note="match: GSS: Em:B85187"
repeat_region 92101..92152 /note="26 copies 2 mer tt 71% conserved"
misc_feature 94297..95612 /note="CpG island"
/evidence=not experimental
misc_feature complement(98493..98894) /note="match: GSS: Em:AQ207463"
repeat_region 103719..104155 /note="match: GSS: Em:AQ543924"
repeat_region 104406..104437 /note="MER57B repeat: matches 303..334 of consensus"
repeat_region 107785..107841 /note="L1MB1 repeat: matches 6111..6171 of consensus"
misc_feature 111249..111767 /note="CpG island"
/evidence=not experimental
misc_feature 113362..113450 /note="match: STS: Em:G16146"
misc_feature match: GSS: Em:AQ806873"

ORIGIN
Query Match 20.5%; Score 336; DB 9; Length 115756;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCTCAAGCAGACGCTAGTCTGGGCGCTTCGTGATGTCCTCAGGAGCTGAGCCGC 231
DB 70572 GCTCAAGCCAGACGCTAGTCTGGGCGCTTCGTGATGTCCTCAGGAGCTGAGCCGC 70631

QY 232 CAACTGGAAGACATCTGAGCACAATCTGTGGACAATAACAGGGGGGGGGGGGGGGG 291
DB 70632 CAACTGGAAGACATCTGAGCACAATCTGTGGACAATAACAGGGGGGGGGGGGGGGG 70691

QY 292 GATGGGGCAGAGTGGCGGCTGAACCGGAGATGACAGAGTCCCGGACCTATGTG 351
DB 70692 GATGGGGCAGAGTGGCGGCTGAACCGGAGATGACAGAGTCCCGGACCTATGTG 70751

QY 352 GCAAGGAATGGGAGGCTTCAACCAACTCCAGTAGTCAATGGAGAGGAAGAACCTCCAG 411
DB 70752 GCAAGGAATGGGAGGCTTCAACCAACTCCAGTAGTCAATGGAGAGGAAGAACCTCCAG 70811
```

```
QY 412 GGGATCCAAACACAGAGAGATCCGGCAGATGACAGGTCCGAGACCGAGACCATCGA 471
DB 70812 GGGATCCAAACACAGAGAGATCCGGCAGATGACAGGTCCGAGACCGAGACCATCGA 70871

QY 472 AGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 507
DB 70872 AGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 70907

RESULT 20
LOCUS AX834462 2194 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 1586 from Patent EPI347046.
ACCESSION AX834462
VERSION AX834462.1 GI:39920597
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,K.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuko,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 1586 24-SEP-2003;
RESEARCH Association for Biotechnology (JP)
FEATURES
source
1. 2194
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 13.7%; Score 225; DB 6; Length 2194;
Best Local Similarity 99.4%; Pred. No. 6.8e-117;
Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGAACCAAGACAAAGACGGGGCTGCCAACAATCCATCCAAAGAGAGAGAGAGAG 60
DB 66 ATGAAGAACCAAGACAAAGACGGGGCTGCCAACAATCCATCCAAAGAGAGAGAGAGAG 125

QY 61 GGACAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 126 GGACAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 185

QY 121 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 186 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 245

QY 181 AGAACGGCTCAGTCTGGGGCGCTTCGTGATGTCCTCAGGAGCTGAGCCGCAACTGGAA 240
DB 246 AGAACGGCTCAGTCTGGGGCGCTTCGTGATGTCCTCAGGAGCTGAGCCGCAACTGGAA 305

QY 241 GACATATCTAGCACAATCTGTGTGGACAATACAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 306 GACATATCTAGCACAATCTGTGTGGACAATACAGAGAGAGAGAGAGAGAGAGAGAGAG 365

QY 301 CAGGTGAGCGCGCTGAACCGCAAGAT 327
DB 366 CAGGTGAGCGCGCTGAACCGCAAGAT 392

RESULT 21
LOCUS AK096939 2194 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ39620 fis, clone SMINT2001114.
ACCESSION AK096939
VERSION AK096939.1 GI:21756551
KEYWORDS clligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
 Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S., Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, T., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuna, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE
 NEDO human cDNA sequencing project
JOURNAL
 Unpublished
AUTHORS
 Isogai, T. and Yamamoto, J.
TITLE
 Direct Submission
JOURNAL
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.
FEATURES
 source
 1..2194
 /organism="Homo sapiens"
 /mol_type="cDNA"
 /db_xref="taxon:9606"
 /clone="SMINT2001114"
 /tissue type="small intestine"
 /clone_lib="SMINT2"
 /note="cloning vector: pME18SFL3"

ORIGIN
 Query Match 13.7%; Score 225; DB 9; Length 2194;
 Best Local Similarity 99.4%; Pred. No. 6.8e-117;
 Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGACCAAGCAAAAGAACCGGGCTGCCAACAATCAATCCAAAGACGCCA 60
 |||||
 Db 66 ATGAAGACCAAGCAAAAGAACCGGGCTGCCAACAATCAATCCAAAGACGCCA 125
 |||||

QY 61 GGACACCGGAGCAGAGCCCGGAGGAGCCAGGAGCGCCAGGCGGCTCTGCA 120
 |||||
 Db 126 GGACACCGGAGCAGAGCCCGGAGGAGCCAGGAGCGCCAGGCGGCTCTGCA 185
 |||||

QY 121 GTAGAAGCAGAGGTCCTCGGAGCAGCCAGGTCCTTCGGAAGCCGGAGGTCCTCAAGCC 180
 |||||
 Db 186 GTAGAAGCAGAGGTCCTCGGAGCAGCCAGGTCCTTCGGAAGCCGGAGGTCCTCAAGCC 245
 |||||

QY 181 AGAACCGCTCAGTCTGGGCGCTTCTGTGATGTCCTGAGGAGCTCAGCGCCCACTGGAA 240
 |||||
 Db 246 AGAACCGCTCAGTCTGGGCGCTTCTGTGATGTCCTGAGGAGCTCAGCGCCCACTGGAA 305
 |||||

QY 241 GACATACTGAGCACACTGTGTGGACATAACAGGGGGGGCCCGGAGGATCGGGCA 300
 |||||
 Db 306 GACATACTGAGCACACTGTGTGGACATAACAGGGGGGGCCCGGAGGATCGGGCA 365
 |||||

QY 301 CAGGCTGAGCGGCTGAACCCGAAGT 327
 |||||
 Db 366 CAGGCTGAGCGGCTGAACCCGAAGT 392
 |||||

RESULT 22
 AR199533
LOCUS
 DEFINITION
 Sequence 14 from patent US 6355451.
ACCESSION
 AR199533
VERSION
 AR199533.1 GI:20249607

KEYWORDS
 Unknown.
SOURCE
 ORGANISM
 Unknown.
REFERENCE
 1 (bases 1 to 4722)
AUTHORS
 Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE
 Low density lipoprotein binding proteins and their use in diagnosing and treating atherosclerosis
JOURNAL
 Patent: US 6355451-A 14 12-MAR-2002;
FEATURES
 Location/Qualifiers
 1..4722
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 3.6%; Score 59; DB 6; Length 4722;
 Best Local Similarity 100.0%; Pred. No. 1.9e-21;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGAGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1454
 |||||
 Db 1456 CGGCTGAGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1514
 |||||

RESULT 23
 AR374685
LOCUS
 AR374685 4722 bp DNA linear PAT 18-DEC-2003
DEFINITION
 Sequence 14 from patent US 6605588.
ACCESSION
 AR374685
VERSION
 AR374685.1 GI:40077500
KEYWORDS
 Unknown.
ORGANISM
 Unknown.
REFERENCE
 1 (bases 1 to 4722)
AUTHORS
 Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE
 Low density lipoprotein binding proteins and their use in diagnosing and treating atherosclerosis
JOURNAL
 Patent: US 6605588-A 14 12-AUG-2003;
FEATURES
 Location/Qualifiers
 1..4722
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 3.6%; Score 59; DB 6; Length 4722;
 Best Local Similarity 100.0%; Pred. No. 1.9e-21;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGAGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1454
 |||||
 Db 1456 CGGCTGAGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAGAGGGT 1514
 |||||

RESULT 24
 AR409320
LOCUS
 AR409320 4722 bp DNA linear PAT 18-DEC-2003
DEFINITION
 Sequence 14 from patent US 6632923.
ACCESSION
 AR409320
VERSION
 AR409320.1 GI:40160108
KEYWORDS
 Unknown.
ORGANISM
 Unknown.
REFERENCE
 1 (bases 1 to 4722)
AUTHORS
 Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE
 Low density lipoprotein binding proteins and their use in diagnosing and treating atherosclerosis
JOURNAL
 Patent: US 6632923-A 14 14-OCT-2003;
FEATURES
 Location/Qualifiers
 1..4722
 /organism="unknown"
 /mol_type="genomic DNA"

```

ORIGIN
Query Match          3.6%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1396 CGGCTGGAGAACTGTGCCGGGCACCTGCACAGAGCGCAATCACTGAACAGAGGGT 1454
        |||||
DB       1456 CGGCTGGAGAACTGTGCCGGGCACCTGCACAGAGCGCAATCACTGAACAGAGGGT 1514

RESULT 25
AX239571 LOCUS              4722 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 14 from Patent WO0164874.
ACCESSION AX239571
VERSION   AX239571.1 GI:15797253
KEYWORDS
SOURCE    Oryctolagus cuniculus (rabbit)
ORGANISM  Oryctolagus cuniculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE
AUTHORS   Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE     Low density lipoprotein binding proteins and their use in diagnosing
          and treating atherosclerosis
JOURNAL   Patent: WO 0164874-A 14 07-SEP-2001;
          Boston Heart Foundation, Inc. (US)
FEATURES
source    1..4722
           /organism="Oryctolagus cuniculus"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9986"
CDS       61..1734
           /note="unnamed protein product"
           /codon_start=1
           /protein_id="CAC88294.1"
           /db_xref="GI:15797254"
           /db_xref="RETRMBL:CAC88294"
           /translations="MKNDKNGAAKQPNPKSPGOPEAGAQAQGPGRPAPARAE
           GASSAPGPGEQAQTAQPGALCDVSELSROLEPILSTVCVDNNGAGGEDGVQGE
           PPPEDAEKSRVAIVANGPEPTPVNGEKEKTSKAEPETEIRTSDEVGDRDHRPQ
           EKXKAGLGKEITLLMOTLNTLSTPEEKLAALCKKYAELEHRSNQKKLLQKQS
           QLVQEDHLRGHEHSKAILAPSKLESICRELQRHNSLKEEGVQVQAREEEKRKVEVTH
           FMQLNDIOLIQMEQHNRNSKLQEENNELAERLKLIEQVELFEEDIKVFKHKDLQY
           OLVDKLIQQAQELKEAESRHOEKDFLLKEAVEASORMCELMEQQETHLKQQLALYTE
           KPFEONTLSKSSEVTTFQOEEMKTKIKLEKETTWYSRWSSNKALLEMAEK
           TLDEKLEGLOVKIQSLKLCRALOTERNLDNKRVDLGNAGGQPYSDSGPPRRPEPA
           TTSKEQGVGEGAQVFNPSPRATDASCAGAPSTEASQTGPQEFYTATA"
```

Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
Novel low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
Patent: JP 2001506983-A 5 29-MAY-2001;
BOSTON HEART FOUNDATION INC
PN JP 2001506983-A/5
PD 29-MAY-2001
PF 26-NOV-1997 JP 1998524870
PF 27-NOV-1996 US 60/031930, 03-JUN-1997 US 60/048547 PI
ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC
A61K39/04, A61K38/17, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC
C07H21/00,
PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
1. .4722
/organism="Aequorea victoria"
/mol_type="genomic DNA"
/db_xref="taxon:6100"

Query Match 3.6%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGCGTGGAGAGCTGTGCGGGGCACTGCGAGACAGAGCGCAATGACCTCAACAAGAGGT 1454
|||||
DB 1456 CGCGTGGAGAGCTGTGCGGGGCACTGCGAGACAGAGCGCAATGACCTCAACAAGAGGT 1514
|||||

RESULT 27
AC111774
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 245095)
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayogaji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Bryant, N., Buay, C., Burch, P., Burrell, K., Blyth, P., Brown, E.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Calderon, M.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgieva, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
Guarinate, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hult, S., Hume, J., Hume, J., Idiebird, D., Jackson, A.,
Karpachy, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensushewa, L., Louisedge, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, P.,

[illegible]

Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, K., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Pastorknak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plazzer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, J., Umani, K., Valas, R., Vera, V., Villagana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 245095)
Worley, K.C.

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 245095)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 9, 2002 this sequence version replaced gi:23321701. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOBQ
Center clone name: CH230-171G2
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 153773 bases at least Q40
Consensus quality: 160170 bases at least Q30
Consensus quality: 164659 bases at least Q20
Estimated insert size: 160182; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

*** NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
*** NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
*** This record will be updated with the finished sequence

* as soon as it is available and the accession number will * be preserved.

* 1 69762: contig of 69762 bp in length
* 69763: gap of unknown length
* 69863: contig of 3827 bp in length
* 73683: gap of unknown length
* 73789: contig of 14658 bp in length
* 88447: gap of unknown length
* 88448: contig of 36066 bp in length
* 124613: contig of unknown length
* 124713: gap of unknown length
* 179270: contig of 54557 bp in length
* 179370: gap of unknown length
* 179371: contig of 45126 bp in length
* 224496: gap of unknown length
* 224497: contig of 3540 bp in length
* 228136: contig of unknown length
* 228137: gap of unknown length
* 228237: contig of 11229 bp in length
* 239466: gap of unknown length
* 239566: contig of 4277 bp in length
* 243843: gap of unknown length
* 243943: contig of 1153 bp in length.
* 243943 Location/Qualifiers

source
1..245095

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-171G2"
88548..89924
/note="wgs contig"
91432..93819
/note="wgs contig"
118848..118080
/note="wgs contig"
124714..126005
/note="wgs contig"
228237..229783
/note="wgs contig"
239566..240945
/note="wgs contig"

ORIGIN

Query Match 3.0%; Score 49; DB 2; Length 245095;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1417 GCATGACAGACAGGCGCAATGACCTGAAACAGAGGTCAGGACCTGA 1465
DB 145976 GCATGACAGACAGGCGCAATGACCTGAAACAGAGGTCAGGACCTGA 146024

RESULT 28

BC060227
LOCUS BC060227 4702 bp mRNA linear ROD 02-DEC-2003
DEFINITION Mus musculus cDNA clone MGC:77972 IMAGE:30357541, complete cds.
ACCESSION BC060227
VERSION BC060227.1 GI:38614270
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 4702)
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932

2 (bases 1 to 4702)

Strausberg, R. Direct Submission Submitted (20-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb-re@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0.

FEATURES

source

1. .4702
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:77972 IMAGE:30357541"
/tissue_type="Brain"
/clone_lib="NIH BMAP_GH0"
/lab_host="DH10B"
/note="Vector: PYX-ASC"
220. .1884
/codon_start=1
/product="Unknown (protein for MGC:77972)"
/protein_id="AAH60227.1"
/db_xref="GI:38614271"

CDS

/translation="MKNQKNGPKAGNSKSGPGBAGPEGAHGRPROTPAGAAE
GTSQAPKTEGARAKAAGPGLCDVSELSQLEDILSTVCVNNQGGPAEEGAQGE
PIEPDEKSRITAAANGPEPGIPVNGEKETSGEGTEIRASDEVGRDHRPQ
EKKAKGLGEITLMTQTLNTLSTPEELAAKCKYAELEHRNSQKMLQKKQS
QVQEKHLRGSKAVLSARLSLELRELQRNSLKBEGVQVAREBEKREKVTSH
FOVNDILOQEQHNEHNSKLRQENMELAKELIQYELREHIDKVFQKDLQO
QVADAKLOAQEMLEAEERHOREKEFLKAEVQSMCELMKQOETHLKLQALYTE
KPEEPNTLSKSEVPTTFKQEMKWTKKIKLEKETMYRSRWSSNKALLEMAEEK
TVRDELEGLQVQLRLEKLCRALQTERNDLAKRVQDLTAGGITDGSRRPEATAS
KKGQVESPGAQPASSPRATDAPCCSGAPSTGTAGTGPGEPPTATA"

ORIGIN

Query Match 2.9%; Score 47; DB 10; Length 4702;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACCTGAATGACATTCAGTCGAGTGAACAGCA 872

Db 1045 CACTTCAGGTGACCTGAATGACATTCAGTCGAGTGAACAGCA 1091

RESULT 29
AL671759/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL671759 141790 bp DNA linear ROD 12-JUL-2002
Mouse DNA sequence from clone RP23-151A15 on chromosome 4, complete
sequence.
AL671759
AL671759.10 GI:22138725
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 141790)
Corby, N.
Direct Submission
Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquersanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 8, 2002 this sequence version replaced gi:21732087.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquersanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-151A15 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6.

FEATURES

source

1. .141790
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clones="RP23-151A15"
/clone_lib="RPCI-23"

ORIGIN

Query Match 2.9%; Score 47; DB 10; Length 141790;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACCTGAATGACATTCAGTCGAGTGAACAGCA 872

Db 62204 CACTTCAGGTGACCTGAATGACATTCAGTCGAGTGAACAGCA 62158

RESULT 30

AC108815

LOCUS

DEFINITION

ACCESSION

VERSION

COMMENT

AC108815 181547 bp DNA linear HTG 08-MAR-2003
Mus musculus clone RP23-2390.3, WORKING DRAFT SEQUENCE, 5 unordered
pieces.
AC108815
AC108815.4 GI:28882391

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.
Mus musculus
Mus musculus (house mouse)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 181547)

AUTHORS

Birren,B., Nusbaum,C. and Lander,E.

JOURNAL

Mus musculus, clone RP23-239013

REFERENCE

2 (bases 1 to 181547)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LeRoque,K., Lamazares,R.,
Landers,T., LeRoque,K., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Melgrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 181547)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepe,Y.,
Collymore,A., Cooke,A., Cooke,P., Cortum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (08-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Mar 8, 2003 this sequence version replaced gi:28195423.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 180713 bases at least Q40

Consensus quality: 180940 bases at least Q30

Consensus quality: 181059 bases at least Q20

Insert size: 172000; agarose-fp

Insert size: 181147; sum-of-contigs

Quality coverage: 8.0 in Q20 bases; agarose-fp

Quality coverage: 7.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 15598: contig of 15598 bp in length
15698: gap of 100 bp
15699 17443: contig of 1745 bp in length
17444 17543: gap of 100 bp
17544 39808: contig of 2265 bp in length
39809 39908: gap of 100 bp
39909 76850: contig of 36942 bp in length
76851 76951: gap of 100 bp
76951 181547: contig of 104597 bp in length.

FEATURES

Location/Qualifiers

source

1..181547
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-239013"
/clone_lib="RPC1-23 Female Mouse BAC"

misc_feature

1..15598
/note="assembly_fragment"
clone_end:SP6
vector_side:left
15699..17443
/note="assembly_fragment"
17544..39808
/note="assembly_fragment"
39909..76850
/note="assembly_fragment"
76951..181547
/note="assembly_fragment"
clone_end:T7
vector_side:right

misc_feature

15699..17443

misc_feature

17544..39808

misc_feature

39909..76850

misc_feature

76951..181547

ORIGIN

Query Match 2.7%; Score 44; DB 2; Length 181547;

Best Local Similarity 100.0%; Pred. No. 9.8e-13;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

544 CTGAGTACCCAGAGGAGAGCTGGCTCTCTGCAAGAAGTA 587

112671 CTGAGTACCCAGAGGAGAGCTGGCTCTCTGCAAGAAGTA 112714

RESULT 31

AR090189/c

LOCUS

DEFINITION

SEQUENCE 309 from patent US 5994076.

ACCESSION

AR090189

VERSION

AR090189.1

KEYWORDS

GI:10016944

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 30)

ChenChik,A. Jakhadze,G. and Bibilashvili,R.

TITLE

Methods of assaying differential expression

JOURNAL

Patent: US 5994076-A 309 30-NOV-1999;

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L20929

Center clone name: 239 O.13

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

```
FEATURES
  source
    Location/Qualifiers
      1..30
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
  Query Match
    Best Local Similarity 1.8%; Score 30; DB 6; Length 30;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1211 GCAGCGAGGTATTCACCACTTCAAGCAGG 1240
  Db 30 GCAGCGAGGTATTCACCACTTCAAGCAGG 1
RESULT 32
  LOCUS
    AR090190
    Sequence 310 from patent US 5994076.
    ACCESSION
      AR090190
    VERSION
      AR090190.1 GI:10016945
    KEYWORDS
      Unknown.
    ORGANISM
      Unknown.
    Unclassified.
  REFERENCE
    1 (bases 1 to 30)
    Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
    TITLE
      Methods of assaying differential expression
    JOURNAL
      Patent: US 5994076-A 310 30-NOV-1999;
    FEATURES
      Location/Qualifiers
        1..30
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match
    Best Local Similarity 1.8%; Score 30; DB 6; Length 30;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 860 AGATGGAACAGCACATGAGCGCAACTCCA 889
  Db 1 AGATGGAACAGCACATGAGCGCAACTCCA 30
RESULT 33
  LOCUS
    AR197224/c
    Sequence 309 from patent US 6352829.
    ACCESSION
      AR197224
    VERSION
      AR197224.1 GI:20247073
    KEYWORDS
      Unknown.
    ORGANISM
      Unknown.
    Unclassified.
  REFERENCE
    1 (bases 1 to 30)
    Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
    TITLE
      Methods of assaying differential expression
    JOURNAL
      Patent: US 6352829-A 309 05-MAR-2002;
    FEATURES
      Location/Qualifiers
        1..30
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match
    Best Local Similarity 1.8%; Score 30; DB 6; Length 30;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 860 AGATGGAACAGCACATGAGCGCAACTCCA 889
  Db 1 AGATGGAACAGCACATGAGCGCAACTCCA 30
RESULT 34
  LOCUS
    AR197225
    Sequence 310 from patent US 6352829.
    ACCESSION
      AR197225
    VERSION
      AR197225.1 GI:20247074
    KEYWORDS
      Unknown.
    ORGANISM
      Unknown.
    Unclassified.
  REFERENCE
    1 (bases 1 to 30)
    Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
    TITLE
      Methods of assaying differential expression
    JOURNAL
      Patent: US 6352829-A 310 05-MAR-2002;
    FEATURES
      Location/Qualifiers
        1..30
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match
    Best Local Similarity 1.8%; Score 30; DB 6; Length 30;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 860 AGATGGAACAGCACATGAGCGCAACTCCA 889
  Db 1 AGATGGAACAGCACATGAGCGCAACTCCA 30
RESULT 35
  LOCUS
    AR259378/c
    Sequence 309 from patent US 6489455.
    ACCESSION
      AR259378
    VERSION
      AR259378.1 GI:27309889
    KEYWORDS
      Unknown.
    ORGANISM
      Unknown.
    Unclassified.
  REFERENCE
    1 (bases 1 to 30)
    Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
    TITLE
      Methods of assaying differential expression
    JOURNAL
      Patent: US 6489455-A 309 03-DEC-2002;
    FEATURES
      Location/Qualifiers
        1..30
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
  Query Match
    Best Local Similarity 1.8%; Score 30; DB 6; Length 30;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1211 GCAGCGAGGTATTCACCACTTCAAGCAGG 1240
  Db 30 GCAGCGAGGTATTCACCACTTCAAGCAGG 1
RESULT 36
  LOCUS
    AR259379
    Sequence 310 from patent US 6489455.
    ACCESSION
      AR259379
    VERSION
      AR259379.1 GI:27309890
    KEYWORDS
      Unknown.
    ORGANISM
      Unknown.
    Unclassified.
  REFERENCE
    1 (bases 1 to 30)
    Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
    TITLE
      Methods of assaying differential expression
    JOURNAL
      Patent: US 6489455-A 310 03-DEC-2002;
    FEATURES
      Location/Qualifiers
```

source 1..30
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 1.8%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 860 AGATGGAACACGCAATGAGCGCAACTCCA 889
|||||
Db 1 AGATGGAACACGCAATGAGCGCAACTCCA 30

RESULT 37

LOCUS BX004876/c
DEFINITION Zebrafish DNA sequence from clone DKEY-11M19, complete sequence.
ACCESSION BX004876
VERSION BX004876.7 GI:30268429
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 174893)
Mashreghi-Mohammadi, M.
Direct Submission
Submitted (01-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 30, 2003 this sequence version replaced gi:29786472.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
beginning 'Dr' were identified by the Recon repeat discovery system
(Zhiron Bao and Sean Eddy, submitted), and those beginning 'drr'
were identified by Rick Waterman (Stephen Johnson lab, WashU). For
further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-11M19
is from a zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES

source

1..174893
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-11M19"
/clone_lib="Daniokey"

ORIGIN

Query Match 1.8%; Score 30; DB 5; Length 174893;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 888 CAAGCTGGCCCAAGACATGGAGCTGGC 917
|||||
Db 37933 CAAGCTGGCCCAAGACATGGAGCTGGC 37904

RESULT 38

LOCUS G94938/c
DEFINITION S208P6370PB7.T0 129S1/SvImJ Mus musculus STS genomic, sequence
tagged site.
ACCESSION G94938
VERSION G94938.1 GI:22745694
KEYWORDS STS.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Made, C.
Polymorphism Structure in the Mouse
Unpublished (2002)
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 609
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGS03 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated.
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..609
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S1/SvImJ"
/db_xref="taxon:10090"
/map="4 18-549 127132950-127133480"
/clone_lib="129S1/SvImJ"
<1..5609

STIS

ORIGIN

Query Match 1.8%; Score 29; DB 11; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1258 AAGAAGATCAAGAGCTGGAGAGAAAC 1286

Db 29 AAGAAGATCAAGAGCTGGAGAGAAAC 1

RESULT 39

LOCUS AC107814
DEFINITION Mus musculus clone RP23-95A15, WORKING DRAFT SEQUENCE, 7 ordered
pieces.
ACCESSION AC107814
VERSION AC107814.4 GI:22381962

ORIGIN

KEYWORDS SOURCE ORGANISM

HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 183681)

AUTHORS

Birren,B., Nusbaum,C. and Lander,E.

TITLE

Mus musculus, clone RP23-95A15

REFERENCE

2 (bases 1 to 183681)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Lander,E., Lechoczky,J., Levine,R., Liu,G., MacLean,C.,
MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome

REFERENCE

3 (bases 1 to 183681)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., MacDonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 21, 2002 this sequence version replaced gi:20163114.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L20543

Center clone name: 95_A_15

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 182388 bases at least Q40
Consensus quality: 182936 bases at least Q30
Consensus quality: 183040 bases at least Q20
Insert size: 186000; agarose-fp
Insert size: 183081; sum-of-contigs
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 12.3 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 61449: contig of 61449 bp in length
* 61450 61450: gap of 100 bp
* 61451 61451: contig of 1924 bp in length
* 61452 61452: gap of 100 bp
* 61453 61453: gap of 100 bp
* 61454 61454: contig of 5931 bp in length
* 61455 61455: gap of 100 bp
* 61456 61456: contig of 22732 bp in length
* 61457 61457: gap of 100 bp
* 61458 61458: contig of 26408 bp in length
* 61459 61459: gap of 100 bp
* 61460 61460: contig of 64034 bp in length
* 61461 61461: gap of 100 bp
* 61462 61462: contig of 603 bp in length.

FEATURES

source

1..183681
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-95A15"
/clone_lib="RPC1-23 Female Mouse BAC"

misc_feature

1..61449
/note="assembly_fragment"
vector_side:left
clone_end:SP6
61550..63473
/note="assembly_fragment"

misc_feature

63574..69504
/note="assembly_fragment"
69605..92336
/note="assembly_fragment"
92437..118844
/note="assembly_fragment"

misc_feature

118945..182978
/note="assembly_fragment"
183079..183681
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN

Query Match 1.5%; Score 24; DB 2; Length 183681;
Best Local Similarity 100.0%; Pred.No. 0.31;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 CAGAGCAGATGAGCTCCTACAG 642

DB 139483 CAGAGCAGATGAGCTCCTACAG 139506

RESULT 40

AL833794/c

LOCUS

DEFINITION

Mouse DNA sequence from clone RP23-401J24 on chromosome 2, complete

sequence.

AL833794

ACCESSION

AL833794.9

GI:23380949

linear ROD 27-SEP-2002

KEYWORDS SOURCE ORGANISM

HTG.
Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 206549)

REFERENCE

Wood, J.

Direct Submission

Submitted (27-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Sep 30, 2002 this sequence version replaced gi:23337638.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw,
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-401J24 is
from the RP23-401J24 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

FEATURES source

Location/Qualifiers
1..206549
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="2"
/clone="RP23-401J24"
/clone_lib="RP23-401J24"

ORIGIN

Query Match 1..5%; Score 24; DB 10; Length 206549;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 CAGAGCAGATGAGCTCTCTACAG 642

|||||

Db 72113 CAGAGCAGATGAGCTCTCTACAG 72090

RESULT 41

VW1237991

Vitis vinifera mRNA for putative ripening-related protein (grip32
gene). linear PLN 01-APR-2000

LOCUS

DEFINITION

Accession AJ237991

Version GI:7406674

Keywords grip32 gene; ripening-related protein.

Source Vitis vinifera

Organism Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.

REFERENCE

AUTHORS TITLE

JOURNAL

MEDLINE

PIUMED

2 (bases 1 to 629)

REFERENCE

Davies, C.

Direct Submission

TITLE

JOURNAL

Submitted (31-MAR-1999) Davies C., Plant Industry, Horticulture
Unit, CSIRO, PO Box 350, Glen Osmond, South Australia, 5064,
AUSTRALIA

FEATURES

source

1..629

/organism="Vitis vinifera"

/mol_type="mRNA"

/cultivar="Shiraz"

/db_xref="taxon:29760"

/tissue_type="fruit"

/clone_lib="Shiraz berry cDNA library 10 weeks post
flowering in Lambda ZapII"

/dev_stage="ripening (10 weeks post flowering)"

/country="Australia:South Australia"

1..629

/gene="grip32"

12..392

/gene="grip32"

/functions="unknown"

/codon_start=1

/product="putative ripening-related protein"

/protein_id="CA85631.1"

/db_xref="GI:7406675"

/db_xref="GOA:Q9M4H2"

/db_xref="SPTREMBL:Q9M4H2"

/translation="MAGIHKIEETLHMGGRKEBEKRGQGHGKGQGHGG
YKGGQGHSEYKGGQGHGEGYGGQGHGKINEKIQGHGHEHPVGEKKK
KKEKKKEHGHGHDGSHSSSSSDSD"

ORIGIN

Query Match 1..4%; Score 23; DB 8; Length 629;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 GGAGGAGGAGGAGGAGGAGG 814

|||||

Db 68 GGAGGAGGAGGAGGAGGAGG 90

RESULT 42

BC047810

LOCUS

DEFINITION

Accession BC047810

Version BC047810.1

Keywords MGC.

Source Danio rerio (zebrafish)

Organism Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 2967)

REFERENCE

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Grancini, P., Frange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Davies, C. and Robinson, S.P.
Differential screening indicates a dramatic change in mRNA profiles
during grape berry ripening. Cloning and characterization of cDNAs
encoding putative cell wall and stress response proteins
Plant Physiol. 122 (3), 803-812 (2000)

20177861

10712544

2 (bases 1 to 629)

Davies, C.

Direct Submission

Submitted (31-MAR-1999) Davies C., Plant Industry, Horticulture
Unit, CSIRO, PO Box 350, Glen Osmond, South Australia, 5064,
AUSTRALIA

FEATURES

Location/Qualifiers

1..629

/organism="Vitis vinifera"

/mol_type="mRNA"

/cultivar="Shiraz"

/db_xref="taxon:29760"

/tissue_type="fruit"

/clone_lib="Shiraz berry cDNA library 10 weeks post
flowering in Lambda ZapII"

/dev_stage="ripening (10 weeks post flowering)"

/country="Australia:South Australia"

1..629

/gene="grip32"

12..392

/gene="grip32"

/functions="unknown"

/codon_start=1

/product="putative ripening-related protein"

/protein_id="CA85631.1"

/db_xref="GI:7406675"

/db_xref="GOA:Q9M4H2"

/db_xref="SPTREMBL:Q9M4H2"

/translation="MAGIHKIEETLHMGGRKEBEKRGQGHGKGQGHGG
YKGGQGHSEYKGGQGHGEGYGGQGHGKINEKIQGHGHEHPVGEKKK
KKEKKKEHGHGHDGSHSSSSSDSD"

AAIIIAQDDCTTIGTTFKLLDSFEGLLDREVIANDLXKHTDILLHSYARDLKVDADL
FHQYKDRPIVAKNSAAYVWVGLMERIKDPMDRLTLMNKNWLESFLEPQRTY
DHLMEETRYTRAVDAWCAQVAATSEKLNPLLSLIEBTADGIVLGVDPALVR
LURETKYFLLSTQKNADNAEAAEGEVEVVKAPKLSFDPDSAKDLFASADTY
ROOIISALDICTSYNQVORTILAEKPLVQOKLDAVEQALNGLAEALNWKACADITY
KECEMLVKDVLVNTIKNVKATOGILAMWKNLMFERKDGKTYTFDELNDAFNQLI
QOHSIRADAGKEITKLLSSNRVLKVSKAASRWYVDYFSNIVIDGFSAAIISTVR
YLUSQIDPDLAKTESPDLLEIQVELVAFDPIVKPDLGEGGAKPGLRMIKWLQSLF
IEQLMKRLDVGNGEKLLEEDFEVYDPLANQVMVTLANESRCDKFNQKAFMGYVW
KQDLQATLOQITDNGVTLPDGTDRPDLAKFEQIYKKNVASEAFKDTMNGYV
KNVAKPLQALSTWASKWVLFTHYLOEKVANSITELYTFMRTSTLTLKVMGSEV
ERPEYHPDQDPEAAKAAEBEKKALYAIMACMRDIRRTERGTDTMPPLKXTV
TALHTFGIOLSDVTLHQLDAENWETLAKKMLNREQLAPIQQAEEVIERKSDAFN
RVEDRTPFORKAPFVSGGELKLBQVPAYKLLDEFSGSLEGVPSVLGIAESQ
LQEAQDLFELYQIYQORSEELGHLKSLMDVTGTVNFTFRDMPKIDVDLFL
WEETKSLKXDIRMLNKAENVYDHYRMLBEAIAVLTSLPLVQDLHPAMREHWKLLM
QTTGKHFVMDKFCGLGDLALVDDVACSEIVDAKELNIEKQKIEEDTAGLS
LAFSTYQDSVMAVLDVAEALNADNLQNLQKQYVQSNPMFLETYSKQNMNG
RVSALVETQNVQKWNLESIFIGSADIRVQLPEDSKRFDAVNADFOELMRTAPDIT
NVVEACTLQDROERLENQSMLEQCEKALQEVLETKRVAFPRFVSPADLADILSKG
SNPQILRLHQLKCFDNIILNLSFKRBERGPTKIAIHMHSKEGEVEFVEDCSDGPVE
VWQNVVDSMKALQVFEKRAIPTDYDELPTKMIYVISAQNTVVVFFTFQIEINAF
DLEEGNEALKVELDROVQQLADEINKQTSIDRKKLTICTIDVHSDILVOKL
IDERVEDQCFWQSLRYIQSEKTKTCQVNICDAEIAYSYIYIGNCGCLCTPLTDR
CFITLTQAOQLVLGAPAGPAGTKETTKTLARALGIQYVFNCSDDMDYKAMGTY
KGLAOTGAMCQDFENRI PAVLSVCSQVTKVLDARAKKERTFEDADI SLKGTVM
AFTIMNPGVGRALPESLKAIPRPMVMTVPDLALICEIMLAEGFOMSKLSRFEVI
LYKCEDLLSKSHYDWKRAIKTILYVAGGKRAAPSEBKVLLRALDRNLKSLT
ADDTSLFMGLNLDLPKLELLEPRALDPAFDEAAHKAATELGYDQDQLLKLISVRE
LFVVRVSVLLGAAGCTKAVMTLRLAQSSEKTIYQAVNPKAVTRNELGYLHPA
TREWKEGLSVFTFRNMANNKTHQWIVLDGIDDAEMISMTVMDNNKMLTAENR
IPLTPSNRILLEINMHWSPATVPLPAGVGFINADVQWQPVVAVSWIDKLEAAYRPL
LTALFTKYDPCLEHCRNFKTVPLPAPVNOAMTICKILEGILPKETVAGAPPPKPL
LHVHFVAPCAWAGGQWVDVDTQPSKMWSEKDVQPEKGLAVDYVYDQNC
IMVNEBRTVKFYIIGDFTSLFVPTVTRITRYIFELDSLVSNNHYAMFVNGTGTGSA
IMVNEBRTVMTSFTYINMNSLAPALQVILEQPELKKSGVRIGPGSRMRYFV
DNMMPLVDYDTQSSIELLRQMDVYGVYDKVQKLEINCOAACNPNFAGSELNFI
TPMRQHFVTFVAMPNNAETIRANYQIIDGHFSFDDVAVKNNKLVDATCELERN
MNFPLSAVKEFYQVNLRLDSNITQGLTRAIKEYREPVKVARLWHECERFERDRI
NEADMAKDFEVRVATKFFDDCGGVAIEERPLIYASHASMTYTPEDVPVVALSSY
DVLKLTLEKLYNESNAVMIDLAVFOQAMEHVTRARIIDLPRGNAMLVGGSGSKQ
SLARASYICGYEYQIYVSSTYGVINDFKENLLGLYKAGTKGTPTIFLMTDNIQVKE
GLVLYINDLSTYIADLTFTGINDFAVNEVKAAGILDSACNCDFFIDKVRKF
LHVLFCSPVGRDFRIRARQFPALVNTDFMHPGPEALVSVQAFRLVDVPPNVEEV
VRENIAYHAYAHQCSEASERPEKAFREYNTTPKSYLELSLYKMLLQKRDLLR
SKERLENGIDKIAQAAQVTDLQORVLEKQIYVDEKKAQTDDELIVSIGKEKAIYQVAV
EAGREDEKATLQTSVAFQACERDLLEAPEIQAQAAALNSLNKLSKLSFGS
PAAEIYQVAAACLVTCGCKIPKRDWNAKGMADVNSFLSLMNFCDKONVPVUVE
VBEKDIYSNPGFTPDNIKGSACAGLCSWVINIKYFRILYQVAPKRAALAEANKL
DTANKKLVIRDEVKQLQORVALLEQSLMKATEDKNAIAQADRTAKQAQMAERLNG
LSGENTRWGAIEKRLSLEGLVGDVLIASAFVSYPAGPNMQFRKSLVDEKWLPIIE
RQIMPTQGRPLDLDLTDATKAWANEGLFTDPLSVENGAIMSNASRWALMDPDLQ
IKWINKETNGVLVIQQSQPKYIDOVINCINGWPLLIENLPVIDIDAVLDVIGMT
IKKGRNIMKIGDAEVQDSRFLYLTQKLSNPHKPEVAOATLWFCVTEBKLDEO
LLALVVDHERPDLQEAAGLVSLENEYNITLVEENLILFNLANATGILNILELIEG
LEBKTAIEIEKKAQTEIQAIAKEVVRPVARTRGSLTYFLLIDNLNLDVRYHY
SMANFVFLKGMMDTPGDESKVFLAEKLNQEDLDRKVELLVEITTCFVLIGVAVQ
GLFERHKLIVATQCMQILRSRGEHYAKFEYLLGPKVPMGADNPLHDVSVWGSV
QALKELDYQGLLEDLIGSSRWHELEPEDEPLPGDKRMQEFKLLFRALRP
DRUTSANGRVNMLGAKVTYQPYDLSRYQDASGPTPIFVFLSPGVDAVGSVALG
KLGLFTLNGKVASVLSGOQBPAMDRLSAAHKGWVLLQNIHLTIDWITNOLDKK
NEILSGCAKQASFAIYVAILCVFHAALLERKKGFGVNLPGARSIGKWMYPRYTG
DNNEFPFGPANTANHQVLEFIDEVMPPTPLAFGLHPNABIGFKLREAESFCNS
LVQLOPRESSEGGMGAEERAKVLIDEVVDKLPDIDFMDVRSKINPDPPNPFVWVA
IOESERNMLLASMKSLLELDLGLKGLDITPEMERLILKALATDAVPGSWARNLAYPS
LRPLSGMLNLLARHQAQLVDWTAELSTPRAVLSGLFNQSPFLTAQVQATARNNDWPL
DRTVITTEVTKKQPDQIBANSRDGAIFHGLTLEGAEWDDKIGALDSDSKPELFCMPV
ILVRAVTDKAEKMDYKPCVYTYTEARFREEVFAQLKSKHTTEIKWLAGVCLFLDVV

intron

679..757

/gene="ODA4 or SUP1 wt"

LOCUS

AC127323/c

AC127323

208152 bp

DNA

linear

ROD 08-NOV-2003

/number=1
/evidence=experimental
837..903
/gene="ODA4 or SUP1 wt"
/number=2
997..1185
/gene="ODA4 or SUP1 wt"
/number=3
1346..1517
/gene="ODA4 or SUP1 wt"
/number=4
1718..1907
/gene="ODA4 or SUP1 wt"
/number=5
2307..2491
/gene="ODA4 or SUP1 wt"
/number=6
2960..3351
/gene="ODA4 or SUP1 wt"
/number=7
3996..4226
/gene="ODA4 or SUP1 wt"
/number=8
4798..4981
/gene="ODA4 or SUP1 wt"
/number=9
5237..5430
/gene="ODA4 or SUP1 wt"
/number=10
5764..5943
/gene="ODA4 or SUP1 wt"
/number=11
6080..6218
/gene="ODA4 or SUP1 wt"
/number=12
6605..6752
/gene="ODA4 or SUP1 wt"
/number=13
7146..7302
/evidence=experimental
/gene="ODA4 or SUP1 wt"
/number=14
7546..7714
/gene="ODA4 or SUP1 wt"
/number=15
8132..8308
/gene="ODA4 or SUP1 wt"
/number=16
8461..8613
/gene="ODA4 or SUP1 wt"
/number=17
8910..9075
/gene="ODA4 or SUP1 wt"
/number=18
9462..9485
/gene="ODA4 or SUP1 wt"
/note="P-loop"
10203..10371
/gene="ODA4 or SUP1 wt"

Query Match 1.4%; Score 23; DB 8; Length 20258;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 CGAGGAGGAGGAGGAGGCGCAGG 814
|||||
DB 6235 CGAGGAGGAGGAGGAGGCGCAGG 6257
|||||

RESULT 44

AC127323/c

LOCUS


```

repeat_region 27280..27732
/rpt_family="L1"
repeat_region 28110..28784
/rpt_family="ERVK"
repeat_region 28785..28867
/rpt_family="ERVK"
repeat_region 28871..28930
/rpt_family="tRNA-Lys-AAA"
repeat_region 28908..28999
/rpt_family="ERVK"
repeat_region 29270..29591
/rpt_family="MaLR"
repeat_region 29687..29776
/rpt_family="ERVK"
repeat_region 30796..30869
/rpt_family="Alu"
repeat_region 30870..31143
/rpt_family="L1"
repeat_region 31144..31240
/rpt_family="B4"
repeat_region 32556..32694
/rpt_family="L1"
repeat_region 33391..34307
/rpt_family="L1"
repeat_region 34290..34585
/rpt_family="L1"
repeat_region 34600..34974
/rpt_family="L1"
repeat_region 34994..35060
/rpt_family="B4"
repeat_region 35313..36990
/rpt_family="L1"
repeat_region 37563..37626
/rpt_family="Alu"
repeat_region 38960..39183
/rpt_family="ERVK"
repeat_region 39318..39825
/rpt_family="RMER19B"
repeat_region 40215..41233
/rpt_family="L1"
repeat_region 41229..42404
/rpt_family="L1"
repeat_region 42646..42813
/rpt_family="RMER15"
repeat_region 42991..43387
/rpt_family="L1"
repeat_region 44881..44980
/rpt_family="L1"
repeat_region 44968..45523
/rpt_family="L1"
repeat_region 45566..46221
/rpt_family="L1"
repeat_region 46332..46453

Query Match 1.4%; Score 23; DB 10; Length 208152;
Best Local Similarity 100.0%; Pred.No.1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 AGAGGATGTGTGAGCTGATCAAG 1131
Db 124985 AGAGGATGTGTGAGCTGATCAAG 124963
|||||
|||||

RESULT 45
HLIGA1C/H
LOCUS HLIGA1C
DEFINITION H.lar gene (allele-2) for Ig alpha-1 heavy chain constant region.
ACCESSION X53708
VERSION X53708.1 GI:2190364
KEYWORDS direct repeat; hinge region; IgA heavy-chain constant region;
immunoglobulin alpha-1 constant region gene.
SOURCE Hylobates lar (common gibbon)
ORGANISM Hylobates lar

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.

1 Kawamura.S., Tanabe,H., Watanabe,Y., Kurosaki,K., Saitou,N. and Ueda,S.
Evolutionary rate of immunoglobulin alpha noncoding region is greater in hominoids than in Old World monkeys
Mol. Biol. Evol. 8 (6), 743-752 (1991)
92130799
1775062

2 Kawamura.S., Saitou,N. and Ueda,S.
Concerted evolution of the primate immunoglobulin alpha-gene through gene conversion
J. Biol. Chem. 267 (11), 7359-7367 (1992)
92218385
1559979

3 Kawamura.S.
Direct Submission
Submitted (22-JUN-1990) Kawamura S., Dept. of Anthropology, Faculty of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
revised by [5]
4 (bases 1 to 2804)
Kawamura.S.
Direct Submission
Submitted (10-JUN-1997) Kawamura S., Dept. of Anthropology, Faculty of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
On Jun 11, 1997 this sequence version replaced gi:22945.
Related sequence: X53387.

FEATURES
source
1..2804
/organism="Hylobates lar"
/mol_type="genomic DNA"
/db_xref="taxon:9580"
/clone="ch28.HL.Ig alpha.20"
/cell_type="leucocyte"
/clone_lib="genomic"
/gemline
join(<290..595,810..1139,1363..1757)
codon_start=3
product="immunoglobulin alpha-1 heavy chain"
protein_id="CAA37745.1"
/db_xref="GI:2190365"
/translation="SPTSPKVFPLSLCTOPDGNVVAACLVQGFPPFPOEPLSVTWSESGQVATNPPSPQDASGLYTTSSQLTPATECLAGKSVTCHVSHYTNPSODVTPCRVPLPTPPCCQPRLSLHRPALEDLLIGSEANLCTLGLRDASGVTFWTSSGKNRVQGPSEDLCCGYSVSSILPGCAFPHNHGKFTCTAAQSESTQTATLSKSGNTFRPEVHLLPPSEELALNELVLTCLARGFSPEDVLVWLCQSQDELPREKYLTVASQEPSPQGTITFAVTSILRVAEDWKKGDTFSCMVGHEALPFTOKTIDRLAGKPTHVNVVM AEVDGTCY"
290..595
/product="constant region of Ig heavy chain"
/note="CH1 domain"
796..810
/note="hinge region"
810..1139
/rpt_type=DIRECT
/product="constant region of Ig heavy chain"
/note="CH2 domain"
811..819
/note="hinge region"
/rpt_type=DIRECT
820..834
/note="hinge region"
/rpt_type=DIRECT
1363..1757
/product="constant region of Ig heavy chain"
/note="CH3 domain"
1782..1787

C_region
repeat_region
C_region
repeat_region
repeat_region
C_region
polyA_signal

1. *Phragmites australis* (Cav.) Trin. ex Steud.

```

/produced="constant region of Ig alpha-2 heavy chain"
/notes="CH2 domain"
810..824
/notes="hinge region"
/rpt_type=DIRECT
1341..1738
/produced="constant region of Ig alpha-2 heavy chain"
/notes="CH3 domain"
1763..1768

polyA_signal
ORIGIN
Query Match      1.3%; Score 22; DB 9; Length 2981;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1508 GGAGGCCAGAGGGCGCTGGGCG 1529
      |||||
Db 2487 GGAGGCCAGAGGGCGCTGGGCG 2466

RESULT 48
AC127867
LOCUS AC127867 171231 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-246B16, *** SEQUENCING IN PROGRESS ***
AC127867
VERSION AC127867.3 GI:25139796
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 171231)
REFERENCE
AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheva, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwankwelen, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Pannkoeh, C.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindecker, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. B., Smajls, D.,

```

```

Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 171231)
Worley, K. C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 171231)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23915276.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZSR
Center clone name: CH230-246B16
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 165516 bases at least Q40
Consensus quality: 166836 bases at least Q30
Consensus quality: 167566 bases at least Q20
Estimated insert size: 170907; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 171231: contig of 171231 bp in length.
Location/Qualifiers
1. 171231
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-246B16"
1. 2117
/note="wgs_end_extension"
misc_feature
FEATURES
source

```



```

misc_feature      clone_end:Sp6"
                  complement(3183..3763)
                  /note="clone boundary
clone_end:Sp6"
site:
end_sequence:B2167524"
complement(164934..165478)
/note="clone boundary
clone_end:T7"
site:
end_sequence:B2167523"
169979..171231
/note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match      1.3%  Score 22;  DB 2;  Length 171231;
Best Local Similarity 100.0%;  Pred. No. 4.4;
Matches 22;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 612 GAATTCACAGAAGCAGATGAAG 633
Db 137294 GAATTCACAGAAGCAGATGAAG 137315

RESULT 49
AC099192/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-171B15, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC099192
AC099192.8 GI:30522018
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 245826)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davilla, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, V., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheewa, L., Loulseg, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Mitjka, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeleneh, O., Okunnu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Popper, F., Poindexter, A., Popovic, D., Primus, E., Fu, L.-L.,

```

```

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
2 (bases 1 to 245826)
Worley, K.C.
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 245826)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHAX
Center clone name: CH230-171B15
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 225041 bases at least Q40
Consensus quality: 228438 bases at least Q30
Consensus quality: 227536 bases at least Q20
Estimated insert size: 232031; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4395: contig of 4395 bp in length
* 4396 gap of unknown length
* 4496 243516: contig of 239021 bp in length
* 243517 243616: gap of unknown length
* 243617 245826: contig of 2210 bp in length.

```

COMMENT

FEATURES
 source Location/Qualifiers
 1. 245826
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-171B15"
 4496. 5980
 /note="wgs contig"
 misc_feature
 21704. 22885
 /note="wgs contig"
 misc_feature
 24266. 243516
 /note="wgs contig"

ORIGIN
 Query Match 1.3% Score 22; DB 2; Length 245826;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 612 GAATTCACAGAGCAGATGAAG 633
 |||||
 Db 168394 GAATTCACAGAGCAGATGAAG 168373

RESULT 50
 AC097345 262624 bp DNA linear HTG 10-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-200F16, *** SEQUENCING IN PROGRESS
 ,, 8 unordered pieces.
 ACCESSION AC097345
 VERSION AC097345.6 GI:30521027
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 262624)
 Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaraike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, K., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, J., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Derano, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howell, S., Hulyak, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J.,
 Kowis, C., Kraft, C.B., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuewa, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mucic, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwankwelu, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.

Unpublished
 Direct Submission
 2 (bases 1 to 262624)
 Worley, K.C.
 Direct Submission
 Submitted (16-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 262624)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GIDL
 Center clone name: CH230-200F16
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 22847 bases at least Q40
 Consensus quality: 228285 bases at least Q30
 Consensus quality: 231776 bases at least Q20
 Estimated insert size: 236040; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: this is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved
 * 1 97284: contig of 97284 bp in length
 * 97385 97384: gap of unknown length
 * 97385 216537: contig of 119153 bp in length
 * 216538 216537: gap of unknown length
 * 216538 244231: contig of 27594 bp in length
 * 244232 244331: gap of unknown length
 * 244332 256199: contig of 11868 bp in length

* 256200 256299: gap of unknown length
 * 256300 contig of 1206 bp in length
 * 257505 257605: gap of unknown length
 * 257506 contig of 1232 bp in length
 * 258837 258937: gap of unknown length
 * 258938 contig of 1570 bp in length
 * 260507 260607: gap of unknown length
 * 260508 contig of 1570 bp in length
 * 260608 262624: contig of 2017 bp in length.

FEATURES

source Location/Qualifiers

1. 262624
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-200F16"
 44632..84449
 /note="clone_boundary"
 clone_end:T7
 site:ECORI
 end_sequence:BH300075
 97385..98497
 /note="wgs_contig"
 216638..218523
 /note="wgs_contig"

ORIGIN

Query Match 1.3%; Score 22; DB 2; Length 262624;

Best Local Similarity 100.0%; Pred. No. 4.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 619 CAGAAGCAGATGAAGTCTCTAC 640

DB 13912 CAGAAGCAGATGAAGTCTCTAC 13933

RESULT 51

AC130864/c

LOCUS Rattus norvegicus clone CH230-22705, *** SEQUENCING IN PROGRESS
 DEFINITION Rattus norvegicus clone CH230-22705, *** SEQUENCING IN PROGRESS
 ***, 14 unordered pieces.

ACCESSION AC130864.3 GI:30581363

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 328901)

REFERENCE

AUTHORS Muzny, D. Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, N., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C. L., Lebowitz, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loullesed, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, K., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 328901)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 328901)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23321473.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKRN

Center clone name: CH230-22705

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 298739 bases at least Q40

Consensus quality: 303892 bases at least Q30

Consensus quality: 307298 bases at least Q20

Estimated insert size: 337164; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbankdraft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 14 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved

1 274661: contig of 274661 bp in length
 274662 274761: gap of unknown length
 274762 278830: contig of 4069 bp in length
 278831 278930: gap of unknown length
 278931 290431: contig of 11501 bp in length
 290432 290531: gap of unknown length
 290532 301913: contig of 11382 bp in length
 301914 302013: gap of unknown length
 302014 305890: contig of 3877 bp in length
 305891 305990: gap of unknown length
 305991 307354: contig of 1364 bp in length
 307355 307454: gap of unknown length
 307455 309339: contig of 1885 bp in length
 309340 309439: gap of unknown length
 309440 310557: contig of 1117 bp in length
 310558 310656: gap of unknown length
 310657 313033: contig of 2377 bp in length
 313034 313133: gap of unknown length
 313134 314274: contig of 1141 bp in length
 314275 314374: gap of unknown length
 314375 317625: contig of 3251 bp in length
 317626 317725: gap of unknown length
 317726 320761: contig of 3036 bp in length
 320762 320861: gap of unknown length
 320862 325629: contig of 4768 bp in length
 325630 325730: gap of unknown length
 325730 328901: contig of 3172 bp in length.

FEATURES

source
 1. 328901
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-22705"
 misc_feature
 1..1975
 /note="wgs contig"
 misc_feature
 242685..244120
 /note="wgs contig"
 misc_feature
 245561..246990
 /note="wgs contig"
 misc_feature
 247041..248892
 /note="wgs contig"
 misc_feature
 249558..254343
 /note="wgs contig"
 misc_feature
 255090..258203
 /note="wgs contig"
 misc_feature
 259146..262533
 /note="wgs contig"
 misc_feature
 262584..266397
 /note="wgs contig"
 misc_feature
 274762..276658
 /note="wgs contig"
 misc_feature
 276709..278830
 /note="wgs contig"
 misc_feature
 278931..280497
 /note="wgs contig"
 misc_feature
 281470..287314
 /note="wgs contig"
 misc_feature
 288814..290431
 /note="wgs contig"
 misc_feature
 290532..292559
 /note="wgs contig"
 misc_feature
 302014..303864
 /note="wgs contig"

ORIGIN

Query Match 1.3%; Score 22; DB 2; Length 328901;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 CAGAAGCAGATGAAGCTCTAC 640
 Db 52827 CAGAAGCAGATGAAGCTCTAC 52806

RESULT 52

AR199549
 LOCUS AR199549 21 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 42 from patent US 6355451.
 ACCESSION AR199549
 VERSION AR199549.1 GI:20249623
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 21)
 AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
 TITLE Low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: US 6355451-A 42 12-MAR-2002;
 FEATURES
 source
 1..21
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGCTCTGAGGAGCTG 225

Db 1 CGTGATGCTCTGAGGAGCTG 21

RESULT 53

AR374701
 LOCUS AR374701 21 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 42 from patent US 6605588.
 ACCESSION AR374701
 VERSION AR374701.1 GI:40077516
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 21)
 AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
 TITLE Low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: US 6605588-A 42 12-AUG-2003;
 FEATURES
 source
 1..21
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGCTCTGAGGAGCTG 225

Db 1 CGTGATGCTCTGAGGAGCTG 21

RESULT 54

AR409336
 LOCUS AR409336 21 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 42 from patent US 6632923.
 ACCESSION AR409336
 VERSION AR409336.1 GI:40160124
 KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 21)
Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 42 14-OCT-2003;
FEATURES Location/Qualifiers
1..21
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 1.3%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGCTCTGAGGAGCTG 225
|||||
Db 1 CGTGATGCTCTGAGGAGCTG 21
RESULT 55
AX239599
LOCUS AX239599 21 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 42 from Patent WO0164874.
ACCESSION AX239599
VERSION AX239599.1 GI:15797274
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 42 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
1..21
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 1.3%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGCTCTGAGGAGCTG 225
|||||
Db 1 CGTGATGCTCTGAGGAGCTG 21
RESULT 56
BD056463
LOCUS BD056463 21 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis.
ACCESSION BD056463
VERSION BD056463.1 GI:22602069
KEYWORDS JP 2001506983-A/21.
SOURCE Aequorea victoria
ORGANISM Aequorea victoria
Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
Aequoreidae; Aequorea.
REFERENCE
1 (bases 1 to 21)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Novel low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: JP 2001506983-A 21 29-MAY-2001;

BOSTON HEART FOUNDATION INC
PN JP 2001506983-A/21
PD 29-MAY-2001
PF 26-NOV-1997 JP 1998524870
PR 27-NOV-1996 US 60/031930, 03-JUN-1997 US 60/048547 PI
ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC
A61K38/04, A61K38/17, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC
C07H21/00,
PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
1..21
source
/organism="Aequorea victoria"
/mol_type="genomic DNA"
/db_xref="taxon:6100"
ORIGIN
Query Match 1.3%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGCTCTGAGGAGCTG 225
|||||
Db 1 CGTGATGCTCTGAGGAGCTG 21
RESULT 57
AX397366/c
LOCUS AX397366 393 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 1581 from Patent WO0212328.
ACCESSION AX397366
VERSION AX397366.1 GI:21068113
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0212328-A 1581 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
1..393
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 1.3%; Score 21; DB 6; Length 393;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 GCAGAGGTCCTCCGCGACGAGC 147
|||||
Db 90 GCAGAGGTCCTCCGCGACGAGC 70
RESULT 58
AX208583/c
LOCUS AX208583 446 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 423 from Patent WO0157207.
ACCESSION AX208583
VERSION AX208583.1 GI:15423006
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer

JOURNAL Patent: WO 0157207-A 423 09-AUG-2001;
CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

source
1..446
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
|||||
Db 68 GCAGAAGGTCCTCCGGCAGCAGC 48

RESULT 59

AX208299/c
LOCUS AX208299 447 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 139 from Patent WO0157207.
ACCESSION AX208299
VERSION AX208299.1 GI:15422722

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer

JOURNAL

Patent: WO 0157207-A 139 09-AUG-2001;
CORIXA CORPORATION (US)

FEATURES

source
1..447
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 447;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
|||||
Db 68 GCAGAAGGTCCTCCGGCAGCAGC 48

RESULT 60

AX393782/c
LOCUS AX393782 458 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 171 from Patent WO0212329.
ACCESSION AX393782
VERSION AX393782.1 GI:19701749

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Wang, T. and Fan, L.
TITLE Compositions and methods for the therapy and diagnosis of head and neck cancer

JOURNAL

Patent: WO 0212329-A 171 14-FEB-2002;
CORIXA CORPORATION (US)

FEATURES

source
1..458
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
|||||
Db 99 GCAGAAGGTCCTCCGGCAGCAGC 79

RESULT 61

AX197689/c
LOCUS AX197689 491 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 144 from Patent WO0151513.
ACCESSION AX197689
VERSION AX197689.1 GI:15388004

KEYWORDS

source
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Algate, P.A.
TITLE Ovarian tumor-associated sequences
JOURNAL Patent: WO 0151513-A 144 19-JUL-2001;
CORIXA CORPORATION (US)

FEATURES

source
1..491
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 491;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
|||||
Db 68 GCAGAAGGTCCTCCGGCAGCAGC 48

RESULT 62

AX360391/c
LOCUS AX360391 506 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 75 from Patent WO0202623.
ACCESSION AX360391
VERSION AX360391.1 GI:18675891

KEYWORDS

source
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Wang, T., McNeill, P.D., Wantanabe, Y., Carter, D., Henderson, R.A. and Kalos, M.D.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: WO 0202623-A 75 10-JAN-2002;
CORIXA CORPORATION (US)

FEATURES

source
1..506
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 95 GCAGAAGGTCCTCCGGCAGCAGC 75
 |||||

RESULT 63
 AX197992/c
 LOCUS AX197992 536 bp DNA linear PAT 29-AUG-2001
 DEFINITION Sequence 447 from Patent WO0151513.
 ACCESSION AX197992
 VERSION AX197992.1 GI:15388307
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCE
 AUTHORS Algate P.A.
 TITLE Ovarian tumor-associated sequences
 JOURNAL Patent: WO 0151513-A 447 19-JUL-2001;
 CORIXA CORPORATION (US)
 FEATURES
 source
 1. .536
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 536;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 68 GCAGAAGGTCCTCCGGCAGCAGC 48
 |||||

RESULT 64
 AX884231/c
 LOCUS AX884231 539 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 94 from Patent EP1033401.
 ACCESSION AX884231
 VERSION AX884231.1 GI:40039185
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCE
 AUTHORS Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
 TITLE Expressed sequence tags and encoded human proteins
 JOURNAL Patent: EP 1033401-A 94 06-SEP-2000;
 Genet (FR)
 FEATURES
 source
 1. .539
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 41. .538
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE98345.1"
 /db_xref="GI:40039186"
 /translation="MLRISNMKVLLAALIGSVFLLPGPSAABEKKGPKVT
 KYVFLRIGDEVDGVIFGLFGKTPKTVDFEVALATGKFGYKNSFHFVVKDFMI
 QGDFTRGDTGSGSIYGRFFDENFKLHYGPGVSNANAGKHPXSPSSSRQX
 TAWLDG"

sig_peptide 41. .139
 /note="score 9.9 seq SVFFLLPGPSAA/DE"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 539;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 135 GCAGAAGGTCCTCCGGCAGCAGC 115
 |||||

RESULT 65
 BD023841/c
 LOCUS BD023841 539 bp DNA linear PAT 27-AUG-2002
 DEFINITION Sequence tag and encoded human protein.
 ACCESSION BD023841
 VERSION BD023841.1 GI:22565064
 KEYWORDS JP 2001269182-A/87.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 539)
 REFERENCE
 AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
 TITLE Sequence tag and encoded human protein
 JOURNAL Patent: JP 2001269182-A 87 02-OCT-2001;
 GENSET
 COMMENT
 OS Homo sapiens (human)
 PN JP 2001269182-A/87
 PD 02-OCT-2001
 PF 24-FEB-2000 JP 2000118773
 PR 26-FEB-1999 US 60/122487
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
 PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
 C12N5/10,
 PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
 G06F15/40
 CC score 9.9
 CC seq SVFFLLPGPSAA/DE
 FH Key Location/Qualifiers
 FT CDS 41. .538
 FT sig_peptide 41. .139.
 Location/Qualifiers
 source
 1. .539
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 539;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGGCAGCAGC 147
 |||||
 Db 135 GCAGAAGGTCCTCCGGCAGCAGC 115
 |||||

RESULT 66
 AX256016/c
 LOCUS AX256016 608 bp DNA linear PAT 10-OCT-2001
 DEFINITION Sequence 167 from Patent WO0170976.
 ACCESSION AX256016
 VERSION AX256016.1 GI:16075056
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCE
 AUTHORS Xu, J., Pyle, R.A. and Stolk, J.A.
 TITLE Compositions and methods for the therapy and diagnosis of ovarian
 and endometrial cancer
 JOURNAL Patent: WO 0170976-A 167 27-SEP-2001;

```

CORIXA CORPORATION (US)
FEATURES
  source
    Location/Qualifiers
      1..608
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
  Query Match      1..3%; Score 21; DB 6; Length 608;
  Best Local Similarity 100.0%; Pred. No. 12;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGCAGCAGC 147
    |||||
Db 97 GCAGAGGTCCTCCGCGCAGCAGC 77

RESULT 67
HUMCYCLO/c
LOCUS HUMCYCLO 833 bp mRNA linear PRI 15-FEB-1994
DEFINITION Human cyclophilin mRNA, 3' end.
ACCESSION M60457
VERSION M60457.1 GI:181249
KEYWORDS cyclophilin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 833)
AUTHORS Hasel,K.W., Glass,J.R., Godbout,M. and Sutcliffe,J.G.
TITLE An endoplasmic reticulum-specific cyclophilin
JOURNAL Mol. Cell. Biol. 11 (7), 3484-3491 (1991)
MEDLINE 91260697
PUBMED 1710767
COMMENT Original source text: Homo sapiens thymus cDNA to mRNA.
FEATURES
  source
    Location/Qualifiers
      1..833
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /cell_line="Turkat"
      /cell_type="T-cell"
      /tissue_type="thymus"
      <1..625
      /codon_start=2
      /product="cyclophilin"
      /protein_id="AA35733.1"
      /db_xref="GI:181250"
      /translation="KVLAAALIASVFFLLIPGSAADKKKPKVTKVYFDLRIG
      EDVGRVIFGLFGKTVKTVDFNVALATGKGFYKNSKFRHVKDFMIQGGDFTRG
      GTGKSIYGERFPDENFKLKHYPGWVSMANAGKDTNGSQFFITTVKTAWLDRKHVF
      GKVLGMEVVRKVESTKTSRDKPLKDLVIADCGKIEVKPPAIAKE"
      <1..73
      sig_peptide
      mat_peptide
      74..622
      /product="cyclophilin"
ORIGIN
  Query Match      1..3%; Score 21; DB 9; Length 833;
  Best Local Similarity 100.0%; Pred. No. 12;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGCAGCAGC 147
    |||||
Db 69 GCAGAGGTCCTCCGCGCAGCAGC 49

RESULT 68
AR274927/c
LOCUS AR274927 851 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 64 from patent US 6506607.
ACCESSION AR274927
VERSION AR274927.1 GI:29707477
KEYWORDS

CORIXA CORPORATION (US)
FEATURES
  source
    Location/Qualifiers
      1..851
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
ORIGIN
  Query Match      1..3%; Score 21; DB 6; Length 851;
  Best Local Similarity 100.0%; Pred. No. 12;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGCAGCAGC 147
    |||||
Db 83 GCAGAGGTCCTCCGCGCAGCAGC 63

RESULT 69
HUMCYPBA/c
LOCUS HUMCYPBA 851 bp mRNA linear PRI 02-NOV-1994
DEFINITION Human cyclophilin B (hCypB) mRNA, complete cds.
ACCESSION M60857
VERSION M60857.1 GI:181334
KEYWORDS cyclophilin B; cyclosporin A-binding protein; peptidylprolyl
isomerase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS Price,E.R., Zydowsky,L.D., Jin,M.J., Baker,C.H., McKeon,F.D. and
Walsh,C.T.
TITLE Human cyclophilin B: a second cyclophilin gene encodes a
peptidyl-prolyl isomerase with a signal sequence
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (5), 1903-1907 (1991)
MEDLINE 91156714
PUBMED 2000394
COMMENT Original source text: Human, cDNA to mRNA.
FEATURES
  source
    Location/Qualifiers
      1..851
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="Unassigned"
      /tissue_lib="lambda-GT10"
      1..851
      /gene="PPIB"
      13..639
      /gene="PPIB"
      /codon_start=1
      /product="cyclophilin B"
      /protein_id="AA52150.1"
      /db_xref="GI:181335"
      /db_xref="GDB:G00-127-610"
      /translation="MKVLLAALIASVFFLLIPGSAADKKKPKVTKVYFDLRIG
      GEDVGRVIFGLFGKTVKTVDFNVALATGKGFYKNSKFRHVKDFMIQGGDFTRG
      DGTGKSIYGERFPDENFKLKHYPGWVSMANAGKDTNGSQFFITTVKTAWLDRKHV
      FGKVLGMEVVRKVESTKTSRDKPLKDLVIADCGKIEVKPPAIAKE"
      13..87
      sig_peptide
      88..636
      mat_peptide
      /gene="PPIB"
      /product="cyclophilin B"
      /note="G00-127-610"
ORIGIN

```


Query Match 1.3%; Score 21; DB 9; Length 851;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 DB 83 GCAGAGGTCCTCCGGCAGCAGC 63

RESULT 70
 LOCUS AX409730/c 893 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 2377 from Patent WO0229103.
 ACCESSION AX409730
 VERSION AX409730.1 GI:21442435
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.
 TITLE Novel genetic markers for leukemias
 JOURNAL Patent: WO 03039443-A 2003 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE);
 Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 2377 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source Location/Qualifiers
 1..893
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. M63573"

ORIGIN
 Query Match 1.3%; Score 21; DB 6; Length 893;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 DB 116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 71
 LOCUS AX779845/c 893 bp DNA linear PAT 14-JUL-2003
 DEFINITION Sequence 2002 from Patent WO03039443.
 ACCESSION AX779845
 VERSION AX779845.1 GI:32696839
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.
 TITLE Novel genetic markers for leukemias
 JOURNAL Patent: WO 03039443-A 2002 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE);
 Ludwig-Maximilian-Universitaet Muenchen (DE);
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
 FEATURES
 source Location/Qualifiers
 1..893
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 1.3%; Score 21; DB 6; Length 893;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 DB 127 GCAGAGGTCCTCCGGCAGCAGC 147

DB 116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 72
 LOCUS AX779846/c 893 bp DNA linear PAT 14-JUL-2003
 DEFINITION Sequence 2003 from Patent WO03039443.
 ACCESSION AX779846
 VERSION AX779846.1 GI:32696840
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.
 TITLE Novel genetic markers for leukemias
 JOURNAL Patent: WO 03039443-A 2003 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE);
 Ludwig-Maximilian-Universitaet Muenchen (DE);
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
 FEATURES
 source Location/Qualifiers
 1..893
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 1.3%; Score 21; DB 6; Length 893;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
 DB 116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 73
 LOCUS HUMSCYLP 893 bp mRNA linear PRI 03-AUG-1993
 DEFINITION Human secreted cyclophilin-like protein (SCYLP) mRNA, complete cds.
 ACCESSION M63573
 VERSION M63573.1 GI:337998
 KEYWORDS secreted cyclophilin-like protein.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 893)
 REFERENCE
 AUTHORS Spik, G., Haendler, B., Delmas, O., Mariller, C., Chamoux, M., Maes, P.,
 Tartar, A., Montreuil, J., Stedman, K., Kocher, H. P., Keller, R.,
 Hiestand, P. C. and Movva, N. R.
 TITLE A novel secreted cyclophilin-like protein (SCYLP)
 JOURNAL J. Biol. Chem. 266 (17), 10735-10738 (1991)
 MEDLINE 91250363
 PUBMED 2040592
 COMMENT Original source text: Human, cDNA to mRNA.
 FEATURES
 source Location/Qualifiers
 1..893
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="T-cells"
 1..893
 /gene="SCYLP"
 22..672
 /gene="SCYLP"
 /codon_start=1
 /product="secreted cyclophilin-like protein"
 /protein_id="AAA36601.1"
 /db_xref="GI:337999"

gene
 CDS

/translation="MLRLSRNMKVLAAALTAGSVFFLLLPQSAADKKKGPQVTV
KVYFDLRIGDEDVGRVIFGLFKPTKVDNFVALATGKGFYKNSKFRHVRKDFMI
QGGDFTRGDTGGKSIYGERFPDENFKLHYGPGVSMANAGKNTNGSQFFITVYKTA
WLDGKHVFGKVLGMEVVRKVESTKTSRDKPLKDVIIADCGKIEVEKFFAIKE"

ORIGIN

Query Match 1.3% Score 21; DB 9; Length 893;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGCGACGACC 147
|||||

Db 116 GCAGAAGGTCCTCCGCGACGACC 96
|||||

RESULT 74

BC001125/c

LOCUS

DEFINITION Homo sapiens peptidylprolyl isomerase B (cyclophilin B), mRNA (CDNA clone MGC:2224 IMAGE:2966791), complete cds.

ACCESSION BC001125

VERSION BC001125.2 GI:37589007

KEYWORDS MGC.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 894)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Hopkins,R.F., Zordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,

Schneerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

12477932

2 (bases 1 to 894)

Strausberg,R.

Direct Submission

Submitted (11-DEC-2000) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Oct 8, 2003 this sequence version replaced gi:12654578.

Contact: MGC help desk

Email: gcaps-r@mail.nih.gov

Tissue Procurement: ATCC/DCRD/BTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

Contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kertman, Anuradha

Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 3 Row: m Column: 13

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA 91: 20149505.

FEATURES

source

1..894

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:2224 IMAGE:2966791"

/tissue_type="Skin, melanotic melanoma."

/clone_lib="NIH MGC 20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..894

/gene="PP1B"

/note="synonym: CYPB"

/db_xref="LocusID:5479"

/db_xref="MIM:123841"

19..669

/codon_start=1

/product="peptidylprolyl isomerase B (cyclophilin B)"

/protein_id="AAH01125.1"

/db_xref="GI:12654579"

/translation="MLRLSRNMKVLAAALTAGSVFFLLLPQSAADKKKGPQVTV

KVYFDLRIGDEDVGRVIFGLFKPTKVDNFVALATGKGFYKNSKFRHVRKDFMI

QGGDFTRGDTGGKSIYGERFPDENFKLHYGPGVSMANAGKNTNGSQFFITVYKTA

WLDGKHVFGKVLGMEVVRKVESTKTSRDKPLKDVIIADCGKIEVEKFFAIKE"

151..633

/note="pro_isomerase; Region: Cyclophilin type

peptidyl-prolyl cis-trans isomerase"

/db_xref="CDD:pfam00160"

ORIGIN

Query Match 1.3% Score 21; DB 9; Length 894;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAAGGTCCTCCGCGACGACC 147
|||||

Db 113 GCAGAAGGTCCTCCGCGACGACC 93
|||||

RESULT 75

BC032138/c

LOCUS

DEFINITION

Homo sapiens peptidylprolyl isomerase B (cyclophilin B), mRNA (CDNA

clone MGC:29698 IMAGE:4896656), complete cds.

ACCESSION BC032138

VERSION BC032138.2 GI:40226561

KEYWORDS MGC.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 897)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 897)
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:21619418.
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Submitted (06-JUN-2002) National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Ahter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granate, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: j Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .897
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:29698 IMAGE:4896656"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC_19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .897
/gene="PP1B"
/note="synonym: CYPB"
/db_xref="locusID:5479"
/db_xref="MIM:123841"
24. .674
/gene="PP1B"
/codon_start=1
/product="peptidylprolyl isomerase B (cyclophilin B)"
/protein_id="AAH32138.1"
/db_xref="GI:21619419"
/db_xref="LocusID:5479"
/db_xref="MIM:123841"
/translation="MLRLSRNMKVLAAALIAAGSVFLLPGLPSAADEKKGPKVTY
KYVFLRIGDEVDGVIPLGLGKTVPTVDNFVALATGKGGYKNSKFRHVIKDFMI
QSGDTRGDTGSGKSIYGERPDENFKLKHVPGWVSNAGKDTNGSQFFITVTKTA
WLDGKHVPGKVLGEMVVRKVESTKTSRDKPLKDLVLIADCGKIEVEKPAIAKE"
156. .638
/gene="PP1B"
/note="pro isomerase; Region: Cyclophilin type
peptidyl-prolyl cis-trans isomerase"
/db_xref="CDD:pfam00160"

misc_feature
1.3%; Score 21; DB 9; Length 897;

ORIGIN
Query Match

Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 127 GCAGAGGTCCTCCGCGCAGCAGC 147
Db 118 GCAGAGGTCCTCCGCGCAGCAGC 98
RESULT 76
BC020800/c
LOCUS
DEFINITION
Homo sapiens peptidylprolyl isomerase B (cyclophilin B), mRNA (CDNA
clone MGC:23783 IMAGE:4248828), complete cds.
ACCESSION
BC020800
VERSION
BC020800.1 GI:18088543
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 908)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Datchenko, B., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bogak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Buffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 908)
AUTHORS
Strausberg, R.
Direct Submission
Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) med@paxil.stanford.edu
R. M.
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 35 Row: d Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20149505.
Location/Qualifiers
1. .908
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="MGC:23783 IMAGE:4248828"
/tissue_type="Prostate"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B"
/notes="vector: pDNR-LIB"
1..908
/gene=
/genes="PP1B"
/notes="synonym: CYPB"
/db_xref="LocusID:5479"
/db_xref="MIM:123841"
26..676
/codon_start=1
/product="peptidylprolyl isomerase B (cyclophilin B)"
/protein_id="AAH20800.1"
/db_xref="GI:18088544"
/db_xref="LocusID:5479"
/translation="MLRLSRNMKVLAAALTAGSVFLLPSPSADEKKGPKVTV
KYFDRIGDEDVGRVIFGLFGKIVPTVDNFVALATGEKGYKNSKHFHVIKDPMI
QGDFRGDTGKGSIYGRFPDENFKLKHGPGWVMANAGKDTNGSQFFITVTKTA
WLDGKHFVFGKLEGMVVRKVESTKDSRDKPLKDVIIADCGKIEVEKPFALAKE"
158..640
/misc_feature
/notes="pro isomerase; Region: Cyclophilin type
peptidyl-prolyl cis-trans isomerase"
/db_xref="CDD:pfam00160"

ORIGIN
Query Match 1.3%; Score 21; DB 9; Length 908;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147
Db 120 GCAGAGGTCCTCCGCGACGAC 100

RESULT 77
BC008848/c
LOCUS
DEFINITION
Homo sapiens peptidylprolyl isomerase B (cyclophilin B), mRNA (CDNA
clone MGC:14109 IMAGE:3502055), complete cds.
ACCESSION
BC008848
VERSION
BC008848.2 GI:33874012
KEYWORDS
MGC.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens (human)
REFERENCE
1 (bases 1 to 1019)
AUTHORS
Strausberg,R., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Tschuyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED
12477932
REFERENCE
2 (bases 1 to 1019)

AUTHORS
Strausberg,R.
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:14250757.
Contact: MGC help desk
Email: gcgaps@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgr.nih.gov
Contact: N. Ayala, K. Beckstrom-Sternberg S.M., Benjamin, B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgueon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 20 Row: a Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20149505.

FEATURES
source
1..1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:14109 IMAGE:3502055"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
1..1019
/gene="pp1B"
/notes="synonym: CYPB"
/db_xref="LocusID:5479"
/db_xref="MIM:123841"
141..791
/codon_start=1
/product="peptidylprolyl isomerase B (cyclophilin B)"
/protein_id="AAH08848.1"
/db_xref="GI:14250758"
/db_xref="LocusID:5479"
/translation="MLRLSRNMKVLAAALTAGSVFLLPSPSADEKKGPKVTV
KYFDRIGDEDVGRVIFGLFGKIVPTVDNFVALATGEKGYKNSKHFHVIKDPMI
QGDFRGDTGKGSIYGRFPDENFKLKHGPGWVMANAGKDTNGSQFFITVTKTA
WLDGKHFVFGKLEGMVVRKVESTKDSRDKPLKDVIIADCGKIEVEKPFALAKE"
273..755
/misc_feature
/notes="pro isomerase; Region: Cyclophilin type
peptidyl-prolyl cis-trans isomerase"
/db_xref="CDD:pfam00160"

ORIGIN
Query Match 1.3%; Score 21; DB 9; Length 1019;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGACGAC 147
Db 235 GCAGAGGTCCTCCGCGACGAC 215

RESULT 78

```

AY191780
LOCUS 1060 bp mRNA linear INV 02-MAR-2003
DEFINITION Branchiostoma belcheri hypothetical muscle-derived protein mRNA, complete cds.
ACCESSION AY191780 GI:28629730
VERSION AY191780.1
KEYWORDS Branchiostoma belcheri
SOURCE Branchiostoma belcheri
ORGANISM Branchiostoma belcheri
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
REFERENCE 1 (bases 1 to 1060)
AUTHORS Zhang, H. and Lin, Y.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-2002) Life Science School, Shandong University, Developmental Biology Institute, 27th Shanda Nan Road, Jinan, Shandong 250100, P.R. China
FEATURES
source 1..1060
/organism="Branchiostoma belcheri"
/mol_type="mRNA"
/db_xref="taxon:7741"
CDS 61..855
/codon_start=1
/product="hypothetical muscle-derived protein"
/protein_id="AA045170.1"
/db_xref="GI:28629731"
/translation="MKEENELCGKLSLVEQYKEEHEVEKLIKQDLERQLAEAKM QATMLAEKEKFLTEKQVLKENYHRSQMAIQENMKMQLKLYTERFELHK TLNQSEVNFKEVDMTKMKLEKESFQWRTKWEKSTVLLMAEEKTRDEL IMLRKTCGLQNALRSDTEGGADKTEVLTKEAEQTLLAEFPKQDSKPK PPKKDETKQPNRDTQLEDAETKQPDAGKTEKRPDMLVCRPQ"
ORIGIN
Query Match 1.3%; Score 21; DB 3; Length 1060;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1168 TACACAGAGAGTTTGAGGAG 1188
DB 337 TACACAGAGAGTTTGAGGAG 357
RESULT 79
PPIGACR 2739 bp DNA linear PRI 15-OCT-1997
LOCUS P. pygmaeus gene for Ig alpha heavy chain constant region.
DEFINITION X53704
ACCESSION X53704 GI:38154
KEYWORDS direct repeat; hinge region; IgA constant region; IgA heavy-chain constant region.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1
AUTHORS Kawamura, S., Tanabe, H., Watanabe, Y., Kurosaki, K., Saitou, N. and Ueda, S.
TITLE Evolutionary rate of immunoglobulin alpha noncoding region is greater in hominoids than in Old World monkeys
JOURNAL Mol. Biol. Evol. 8 (6), 743-752 (1991)
MEDLINE 92130799
PUBMED 1775062
REFERENCE 2
AUTHORS Kawamura, S., Tanabe, H., Watanabe, Y., Kurosaki, K., Saitou, N., Omoto, K. and Ueda, S.
TITLE Evolutionary rate of immunoglobulin alpha noncoding region is greater in hominoids than in old world monkeys
JOURNAL Mol. Biol. Evol. 8, 793-793 (1991)
REFERENCE 3
AUTHORS Kawamura, S.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-1990) Kawamura S., Dept. of Anthropology, Faculty

of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
COMMENT Related sequence: X53385.
FEATURES
source 1..2739
/organism="Pongo pygmaeus"
/mol_type="genomic DNA"
/db_xref="taxon:9600"
/clone="ch28.PP.Ig alpha.28"
/clone_lib="genomic"
/germline
join(<291..596,807..1160,1384..>1778)
CDS 291..596
/codon_start=3
/product="immunoglobulin alpha heavy chain"
/protein_id="CAA37741.1"
/db_xref="GI:2160054"
/db_xref="REMBL:CAA37741"
/translation="SPSPKVFPLNFCISIQPDGNVAVACLVQGFPOEPLSVTWSQSG QVTRNPPSPQDTGGTYTSSLTLPATQCDQKSVTCHVKHYTNPSQDVTPCRV PRPTPTSTPCPPSCGQQLSLHRLPALEDLLGSEANLTCTLTGTNAGSAGFTWT PSSGNVAVQRPDPDLGGCYSSVSLPGCDPWNHGQFTTCTAEHPKLTQLTATLSK SGNTRFEVHLPPPSBELNALNELVLTCLARGSPKDLVLRWLQGSQELFPRKYLTV ASRQEPQSGTITYAVTSLIRVAEDWKKGDTFCWVGHEALPLAFTOKTIDRLAGKPT HNVSVVYMAEVDGTCY"
C_region 291..596
/product="constant region of Ig alpha heavy chain"
/note="CH1 domain"
repeat_region 793..807
/note="hinge region"
/rpt_type=DIRECT
C_region 807..1160
/product="constant region of Ig alpha heavy chain"
/note="CH2 domain"
repeat_region 808..816
/note="hinge region"
/rpt_type=DIRECT
repeat_region 817..831
/note="hinge region"
/rpt_type=DIRECT
repeat_region 826..840
/note="hinge region"
/rpt_type=DIRECT
repeat_region 841..855
/note="hinge region"
/rpt_type=DIRECT
C_region 1384..1778
/product="constant region of Ig alpha heavy chain"
/note="CH3 domain"
polyA_signal 1803..1808
ORIGIN
Query Match 1.3%; Score 21; DB 9; Length 2739;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1508 GGAGGCCAGAGGGCGCTGGG 1528
DB 2544 GGAGGCCAGAGGGCGCTGGG 2524
RESULT 80
HSM804617
LOCUS Homo sapiens mRNA; cDNA DKFZp451N249 (from clone DKFZp451N249).
DEFINITION AL833304
ACCESSION AL833304.1 GI:21733938
VERSION
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8821)
AUTHORS Ottenwaelder, B., Obermaier, B., Deutschenbaier, S., Mewes, H.W.,

Weill.B., Amid.C., Osanger,A., Fobo.G., Han,M. and Wiemann,S.
 Direct Submission
 Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
 sequenced by MedGenomix (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp451N249) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/proj/cDNA/.
 FEATURES
 Location/Qualifiers
 source
 1..8821
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp451N249"
 /db_xref="taxon:9606"
 /clone="DKFZp451N249"
 /tissue_type="human skeletal muscle"
 /clone_lib="451 (synonym: hlccl). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 polyA_signal
 8764..8769
 polyA_site
 8790
 ORIGIN
 Query Match. 1.3%; Score 21; DB 9; Length 8821;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1509 GAGCCACAGAGGGCGCTGGGCG 1529
 Db 4046 GAGCCACAGAGGGCGCTGGGCG 4066
 RESULT 81
 AF528099 9708 bp mRNA linear PRI 13-MAR-2003
 LOCUS Homo sapiens transforming acidic coiled coil 2, long isoform
 DEFINITION (TACC2) mRNA, complete cds.
 ACCESSION AF528099
 VERSION AF528099.1 GI:28933421
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 9708)
 Lauffart,B., Gangisetty,O. and Still,I.H.
 AUTHORS Molecular-cloning, genomic structure and interactions of the
 TITLE putative breast tumor suppressor TACC2
 GENOMICS 81 (2), 192-201 (2003)
 JOURNAL
 MEDLINE
 PUBMED 12620397
 REFERENCE 2 (bases 1 to 9708)
 Still,I.H., Lauffart,B. and Gangisetty,O.
 AUTHORS Direct Submission
 TITLE Submitted (10-JUL-2002) Cancer Genetics, Roswell Park Cancer
 JOURNAL Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA
 FEATURES
 Location/Qualifiers
 source
 1..9708
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10q26"
 1..9708
 /gene="TACC2"
 342..9188
 /gene="TACC2"
 /codon_start=1
 gene
 cds
 /product="transforming acidic coiled coil 2, long isoform"
 /protein_id="AA062630.1"
 /db_xref="GI:28933422"
 /translation="MGNENSTSDNQRTLSAQTPRSQPPGNSQNIKRKQDTPGSDPH
 RDASSIGSVGLGFCPTASSASLADCLVSPTEVTPKRPQDQKPGSGSLPFPSPSQ
 EREPHSGMVPFCECPPCGLASAAAPEDPQTPQSPREAPNAPDIAAFAERDS
 STPQETAAVPSAGRERQKEGQSKSFSSGIDQSPGMSVPLRPMKAPLCEGSD
 QPGFSEQKEAAGGPPAESRQVAVQVTPAPAAQQTGESSALVLEKSLPKMAP
 IPQDPAPRASDRERGQSEAPQYLTDDLELRACHLPRSNAGAPAEVNAASQESQ
 QVQAYLPHALLWGLFPAVFEAGSGKEALTDVQSHGTGRTGKPNVVCVA
 AGQEGGLVPSPPELSLTPTTEAPASSLASFPAAQIPLAVEEPPSSSSSSVSKAK
 VSDAAAEVYDAGLVGLRQSDLGSKGHPGCEPAPSPGTPGAKVHEINTSPAVK
 PQVQPPPLPKQESHEVQPCAPPLPKKAPSESARGPGPTGAKVHEINTSPAVK
 EGRSPGDSFGKKEAPEPPDGGDPGNLOQEDSQAFSSKRDPEVGDLSKPSDAES
 RDHPSHAQPPRKGAGHTDGHSDTADASGLPKHKEEDPVLVPVDPGAGEPTV
 PEGAIWEGSLQPKCPTLQSRGLRMESFLTLESEKSPPTPTVAEVAAPKAGRES
 TLEIRKMSGCDGGLLTPDQPRGACDASRQFPHAGVPPPOGENLAADLGLTALL
 DQDQGGFCFGEGWIRGAASEWPLSSSEKHFQPSQAQPTSFIDVLKQAQPPNGK
 ETSFHPGFKDQGADSSQIHVPVEPQEDNNLPTHGQEQALGSELQSLPGLTSDTP
 TSPDMMWESSLTSESLAPTRQKLPALGKRPPEGACDQSGSSVSPPAADLVDF
 SLAENFSRKTCTCTGQPNKSOQALADALEGQHEACORHPGASEAAQCSPLMGL
 SKRWASNTGEAPPCOPDSVALLDVPCLPALAPASPGVPTQDAPETACDTEQEG
 RQVPVAPQKMECRATSDAESPKLLASFPAGEQGEACAAETGSGAGADGPKQA
 PEKFEATLSCGLLQTEHCLTSGEASTSALRESQAEHMASCCQDALLPARELGSP
 RSTWDFTHQAVDFPKELLSGPPEVAAPDTPYLHVDSAQQAQAGDSGVKAVSSAPR
 AGESPCFVGEPPLENAASLKLIFAGSLAPLQPGAAGEIIPAVQASSGSPKARTE
 GPVDSMPCLDRLMLLAGKQATGKAAATAPGAGAKASGGMADGAGTEGSMERG
 EPSODPKQGTSGGVDTSEQIATLTGPDREHIAKIFKPVLCALATPGEKAGRS
 AVGKDLTRPLGPEKLDGPPGVDTLLPAPARLOVEKQKQALAGEAEIFSLALQDPAS
 DXLLGAGLWERNLPGAGVKGWAGVPTILREDERPEGGAAMPGLGQAYSQLERS
 RQELASGLPAAQQLPVERAAAVAPVSHGEEAQAQDRIFSGKHQHTSACDSFH
 GEDGDFAITGVGHVTRPCAPSPQREVLTVPENSEPTLDTLGGERRPVTAGI
 LEMENLNGNSTPAPTPTGQVADTLEKGVAGAGEAGDITLSTAEQACAGDLPE
 AGTTRTFVAGDLVLPSCQDQACSKAPMEGTALHGDSDPARQQAQEQPGPRP
 IPAGDGKVCVSSPPEDETHDKHLAPELHTDRESPPRPGSMLPSVPKQDAPRM
 DKVTSDETRGAECTESSPVADDIIQAPADLESPTLAAASYHGVGVGVSTDLIAQS
 TTPAAHAGLPPSAARIHVSFAPAGDRVEASTPSCDPAKDISRSDSSEAEETPS
 SHFSVAFDEDDPIASSQTYNLDNFNIELVDVTFQLEPRASDANKQEGKYNTRKSTD
 SLPKTSKTLGRSLQASDFDSSGNGNEPAVALADAYSTGSSASNTLTKTKPRP
 PSLKKTQTKPTPTPVKTEQBPDESILVPSGENLASHTKTESATGEGPSFALLEE
 TLEPFAVGAACPLDSEAGVVPVPSGQVQNSPPVGRKTLPLTTPAEAGEVTPS
 DSGQEDSPAGLSVLEFDYSEDKSDMNQENPPPTKIKGKPKVAKMPLRSPKMK
 TPEKLDNTPASPRSPAEPNDFIAKTYTFIDDKWDDNPNPFSSTKWOESPKLPQ
 QSNFNDPTCDESVDPTKSTKTPSSPSKSPASFEIPASMEANGVDGDLNPAKKK
 KTLKLTDTFRVKSFKPSPLSDPESQDPTFAATPTTPIVAVVHATDEKLAVTNOK
 WTCMTVLEADKQDYPQSDLSLTFVNETKFSSTTEELDYNSYEIETMEKIGSLPQD
 DDAPKQALVLMFDTQSPVSKSPVMSSTPCSSGSEFEALVNTAKNQHPVP
 RGLAPNOSHLOVPEKSKQLEAMGLTSPSEAIETAPGSPASADALLSRLAHVPS
 LCGALDLEPDLAEKNPLFAQKLOELEFAMIRIEALKLARQIALASRSHQAKRA
 LCHDVSISKALYSRIGTAVEKPKAGLLQQPDLSALQIARAEIITKRESEWKD
 KYESREVMEMRKIVAEYKTIQMLEDQREKSVSHQVQQLVLEKEQALDAENSV
 EKSADLFRYKMEVLEGRKNEVLEKCAQEVLSRVKKEQVQALVHAEEKLD
 RANAELIAQVQGAQQAQAHQAHLRKEQLRVDALERTLEQKNKEIIBELTKICDELIAK
 MGKS"
 6177..6224
 /gene="TACC2"
 /note="Region: SFP motif"
 7341..7436
 /gene="TACC2"
 /note="Region: SDP repeat 1"
 7437..7523
 /gene="TACC2"
 /note="Region: SDP repeat 2"
 7740..7766
 /gene="TACC2"
 /note="Region: HARDEKLA motif"
 8592..9185
 /gene="TACC2"
 /note="Region: TACC domain"
 1536
 /gene="TACC2"
 /replace="a"
 misc_feature
 6177..6224
 /gene="TACC2"
 /note="Region: SFP motif"
 misc_feature
 7341..7436
 /gene="TACC2"
 /note="Region: SDP repeat 1"
 misc_feature
 7437..7523
 /gene="TACC2"
 /note="Region: SDP repeat 2"
 misc_feature
 7740..7766
 /gene="TACC2"
 /note="Region: HARDEKLA motif"
 misc_feature
 8592..9185
 /gene="TACC2"
 /note="Region: TACC domain"
 variation
 1536
 /gene="TACC2"
 /replace="a"

```

variation      2829      /gene="TACC2"
                /replace="c"
variation      3973      /gene="TACC2"
                /replace="t"
variation      4007      /gene="TACC2"
                /replace="t"
variation      6087      /gene="TACC2"
                /replace="a"
variation      9230      /gene="TACC2"
                /replace="c:a"
variation      9490      /gene="TACC2"
                /replace="c"
polyA_signal   9646.9651  /gene="TACC2"
polyA_site     9671      /gene="TACC2"

ORIGIN
Query Match      1.3%; Score 21; DB 9; Length 9708;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1509 GAGGCCAGAGGGCCCTGGGGC 1529
      |||||
Db 4928 GAGGCCAGAGGGCCCTGGGGC 4948

RESULT 82
AL137003      113802 bp DNA linear PRI 30-SEP-2000
LOCUS
DEFINITION
Human DNA sequence from clone RPI-151F17 on chromosome 6, complete
sequence.
ACCESSION
AL137003
VERSION
AL137003.12 GI:10443369
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 113802)
AUTHORS
Blakey, S.
TITLE
Direct Submission
JOURNAL
Submitted (30-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT
On Oct 1, 2000 this sequence version replaced gi:10432379.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6

FEATURES
source
Location/Qualifiers
1..113802
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RPI-151F17"
/clone_lib="RPI-1"
85317..85658
misc_feature
/note="Single clone region. Assembly consistent with
restriction digest."

ORIGIN
Query Match      1.3%; Score 21; DB 9; Length 113802;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 CTCCTGCAGTAGAAGCAGAG 133
      |||||
Db 20560 CTCCTGCAGTAGAAGCAGAG 20580

RESULT 83
AL357139      128422 bp DNA linear PRI 02-MAY-2001
LOCUS
DEFINITION
Human DNA sequence from clone RP11-793L10 on chromosome 6, complete
sequence.
ACCESSION
AL357139
VERSION
AL357139.11 GI:13990056
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 128422)
AUTHORS
Hammond, S.
TITLE
Direct Submission
JOURNAL
Submitted (02-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT
On May 7, 2001 this sequence version replaced gi:13398769.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at

```

<http://www.sanger.ac.uk/HGP/Chr6>
 RP11-793L10 is from the library RP11-11.3 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBac3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-793L10. It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-793L10 is at 1 in this sequence.
 The true left end of clone RP11-327K21 is at 128323 in this
 sequence.

FEATURES

source	Location/Qualifiers
repeat_region	1..128422
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="6"
repeat_region	/clone="RP11-793L10"
repeat_region	/clone_lib="RP11-11.3"
repeat_region	33..420
repeat_region	/notes="L1MD3 repeat: matches 7093..7521 of consensus"
repeat_region	470..1115
repeat_region	/notes="L1MC4 repeat: matches 6606..7234 of consensus"
repeat_region	1175..2185
repeat_region	/notes="L1M4 repeat: matches 2653..3764 of consensus"
repeat_region	2347..2605
repeat_region	/notes="Charlie4 repeat: matches 2..263 of consensus"
repeat_region	3103..3381
repeat_region	/notes="L1R28 repeat: matches 693..1011 of consensus"
repeat_region	3374..3486
repeat_region	/notes="L1R28 repeat: matches 462..573 of consensus"
repeat_region	3487..4400
repeat_region	/notes="L1P3 repeat: matches 5228..6146 of consensus"
repeat_region	4427..4616
repeat_region	/notes="L1R28 repeat: matches 308..473 of consensus"
repeat_region	4574..4763
repeat_region	/notes="L1M1 repeat: matches 1..203 of consensus"
repeat_region	5515..5985
repeat_region	/notes="L1R1D repeat: matches 27..505 of consensus"
repeat_region	7084..7168
repeat_region	/notes="L1M11 repeat: matches 275..360 of consensus"
repeat_region	7869..8171
repeat_region	/notes="AluX repeat: matches 7..310 of consensus"
repeat_region	11401..11499
repeat_region	/notes="L1R1J repeat: matches 103..199 of consensus"
repeat_region	12331..12382
repeat_region	/notes="26 copies 2 mer ta 76% conserved"
repeat_region	13272..13554
repeat_region	/notes="AluJ repeat: matches 3..282 of consensus"
repeat_region	13718..13998
repeat_region	/notes="L1ME repeat: matches 5256..5540 of consensus"
repeat_region	14389..14735
repeat_region	/notes="L1ME repeat: matches 5592..5991 of consensus"
repeat_region	16355..17252
repeat_region	/notes="L1P12 repeat: matches 5255..6155 of consensus"
repeat_region	17368..17753
repeat_region	/notes="THE1C repeat: matches 1..371 of consensus"
repeat_region	18122..18173
repeat_region	/notes="26 copies 2 mer at 71% conserved"
repeat_region	18736..18884
repeat_region	/notes="L1M4 repeat: matches 4456..4597 of consensus"
repeat_region	19023..19229
repeat_region	/notes="L1ME3 repeat: matches 5775..5981 of consensus"
repeat_region	19679..20188
repeat_region	/notes="L1TLD repeat: matches 1..497 of consensus"
repeat_region	21872..22191
repeat_region	/notes="MER61A repeat: matches 8..353 of consensus"
repeat_region	22199..24950
repeat_region	/notes="MER61-internal repeat: matches 1842..4550 of consensus"
repeat_region	25017..25380
repeat_region	/notes="MER61-internal repeat: matches 3..403 of consensus"
repeat_region	25383..25731
repeat_region	/note="MER61A repeat: matches 3..354 of consensus"
repeat_region	26463..26895
repeat_region	/note="L2 repeat: matches 2267..2694 of consensus"
repeat_region	27124..27371
repeat_region	/note="L1M4 repeat: matches 4119..4374 of consensus"
repeat_region	27369..27621
repeat_region	/note="L1M4 repeat: matches 4539..4817 of consensus"
repeat_region	27627..27780
repeat_region	/note="L1M4 repeat: matches 3930..4097 of consensus"
repeat_region	28028..28288
repeat_region	/note="L1M4 repeat: matches 3388..3658 of consensus"
repeat_region	28294..28535
repeat_region	/note="MER58 repeat: matches 130..2465 of consensus"
repeat_region	28535..28786
repeat_region	/note="Cheshire repeat: matches 108..361 of consensus"
repeat_region	30243..30541
repeat_region	/note="MER46C repeat: matches 19..329 of consensus"
repeat_region	30979..31419
repeat_region	/note="L2 repeat: matches 2022..2460 of consensus"
repeat_region	31441..31514
repeat_region	/note="37 copies 2 mer at 74% conserved"
repeat_region	31627..32110
repeat_region	/note="MER67D repeat: matches 7..510 of consensus"
repeat_region	33443..33698
repeat_region	/note="L2 repeat: matches 1570..1823 of consensus"
repeat_region	34338..34665
repeat_region	/note="L2 repeat: matches 545..867 of consensus"
repeat_region	34886..35026
repeat_region	/note="L2 repeat: matches 2597..2747 of consensus"
repeat_region	35063..35104
repeat_region	/note="21 copies 2 mer aa 78% conserved"
repeat_region	35884..36423
repeat_region	/note="MLTIG repeat: matches 11..512 of consensus"
repeat_region	39579..39684
repeat_region	/note="L1MC4 repeat: matches 7856..7963 of consensus"
repeat_region	42397..42504
repeat_region	/note="54 copies 2 mer ta 84% conserved"
repeat_region	42526..42569
repeat_region	/note="22 copies 2 mer ta 90% conserved"
repeat_region	42821..43046
repeat_region	/note="THE1C repeat: matches 138..369 of consensus"
repeat_region	43052..43651
repeat_region	/note="L1MD3 repeat: matches 6610..7221 of consensus"
repeat_region	43680..45891
repeat_region	/note="TIGER1 repeat: matches 1..2223 of consensus"
repeat_region	45900..46205
repeat_region	/note="AluJ repeat: matches 9..309 of consensus"
repeat_region	46208..46301
repeat_region	/note="TIGER1 repeat: matches 2233..2313 of consensus"
repeat_region	46326..46760
repeat_region	/note="L1P16 repeat: matches 5707..6157 of consensus"
repeat_region	46820..47114
repeat_region	/note="L1MC/D repeat: matches 5222..5403 of consensus"
repeat_region	47887..48203
repeat_region	/note="AluJ repeat: matches 1..312 of consensus"
repeat_region	48456..48805
repeat_region	/note="L1M9 repeat: matches 5607..5974 of consensus"
repeat_region	48889..49206
repeat_region	/note="L1M4 repeat: matches 4591..4916 of consensus"
repeat_region	50394..51317
repeat_region	/note="MLTIB repeat: matches 4..390 of consensus"
repeat_region	52180..52327
repeat_region	/note="L1MB8 repeat: matches 6028..6168 of consensus"
repeat_region	52423..52709
repeat_region	/note="AluX repeat: matches 2..285 of consensus"
repeat_region	52736..53054
repeat_region	/note="L1M1 repeat: matches 148..477 of consensus"
repeat_region	53084..53308
repeat_region	/note="L1M1 repeat: matches 533..767 of consensus"
repeat_region	53441..53536
repeat_region	/note="48 copies 2 mer aa 62% conserved"
repeat_region	53825..54758
repeat_region	/note="L1M4 repeat: matches -37..830 of consensus"


```

repeat_region 57398..57644
/note="MIR repeat: matches 2..262 of consensus"
repeat_region 57890..58009
/note="MIR repeat: matches 86..213 of consensus"
repeat_region 58314..58762
/note="nigger2a repeat: matches 2..434 of consensus"
repeat_region 59308..59839
/note="MER52A repeat: matches 1202..1755 of consensus"
repeat_region 59993..61023
/note="LIPBa repeat: matches -970..1014 of consensus"
repeat_region 60995..61939
/note="LIPBb repeat: matches 41..997 of consensus"
repeat_region 61942..64140
/note="LIPB2 repeat: matches 3716..5928 of consensus"
repeat_region 64119..64465
/note="LIM1 repeat: matches 688..1040 of consensus"

Query Match 1.3%; Score 21; DB 9; Length 128422;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1268 AGAAGCTGGAGAAAGCA 1288
|||||
Db 126868 AGAAGCTGGAGAAAGCA 126888

RESULT 84
AL390741/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-38F19 on chromosome 6, complete
sequence.
ACCESSION AL390741
VERSION AL390741.12 GI:13751970
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 128461).
Chapman, J.
Direct Submission
Submitted (20-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 22, 2001 this sequence version replaced gi:13446471.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EXBL; Sw:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-38F19 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-38F19 It may be shorter because we sequence overlapping

```

sections only once, except for a 100 base overlap.
The true right end of clone RP11-38F19 is at 128461 in this
sequence. The true left end of clone RP11-358115 is at 69922 in
this sequence. The true right end of clone RPI-8B1 is at 100 in
this sequence.

FEATURES

	Location/Qualifiers	Source
repeat_region	232..432	1..128461
	/note="MER20 repeat: matches 1..218 of consensus"	/organism="Homo sapiens"
repeat_region	460..563	/mol_type="Genomic DNA"
	/note="52 copies 2 mer aa 68% conserved"	/db_xref="taxon:9606"
repeat_region	582..620	/chromosome="6"
	/note="HAL1 repeat: matches 158..193 of consensus"	/clone="RP11-38F19"
repeat_region	621..930	/clone_lib="RPCI-11.1"
	/note="AluJb repeat: matches 2..310 of consensus"	
repeat_region	931..1215	
	/note="HAL1 repeat: matches 193..458 of consensus"	
repeat_region	1216..1570	
	/note="THE1B repeat: matches 1..364 of consensus"	
repeat_region	1571..1936	
	/note="HAL1 repeat: matches 458..905 of consensus"	
repeat_region	2602..2915	
	/note="AluJb repeat: matches 1..312 of consensus"	
repeat_region	3035..4478	
	/note="LIPa2 repeat: matches 4703..6146 of consensus"	
repeat_region	4479..5419	
	/note="LIP repeat: matches 2391..3330 of consensus"	
repeat_region	6099..6282	
	/note="MER5A repeat: matches 4..188 of consensus"	
repeat_region	6283..6387	
	/note="MIR repeat: matches 101..199 of consensus"	
repeat_region	6388..6694	
	/note="AluSp repeat: matches 1..306 of consensus"	
repeat_region	6695..6756	
	/note="MIR repeat: matches 199..262 of consensus"	
repeat_region	7094..7194	
	/note="MIR repeat: matches 85..198 of consensus"	
repeat_region	7649..7995	
	/note="LTR16A repeat: matches 25..445 of consensus"	
misc_feature	8789..9346	
	/note="match: GSS: Em:AQ898390"	
misc_feature	8807..9290	
	/note="match: GSS: Em:AQ412103"	
repeat_region	8974..9471	
	/note="L2 repeat: matches 2194..2700 of consensus"	
misc_feature	complement(10027..10127)	
	/note="match: STS: Em:G41857"	
repeat_region	10118..10215	
	/note="LTR32 repeat: matches 326..424 of consensus"	
repeat_region	10218..10993	
	/note="LIM2 repeat: matches 5255..5981 of consensus"	
repeat_region	14273..14405	
	/note="MIR repeat: matches 58..190 of consensus"	
repeat_region	14809..15017	
	/note="L2 repeat: matches 2401..2612 of consensus"	
repeat_region	15599..21698	
	/note="LIPa5 repeat: matches 1..6143 of consensus"	
repeat_region	21820..22211	
	/note="MLR1C repeat: matches 1..426 of consensus"	
repeat_region	22212..22506	
	/note="AluSx repeat: matches 3..297 of consensus"	
repeat_region	22507..22550	
	/note="MLR1C repeat: matches 426..466 of consensus"	
repeat_region	23309..23610	
	/note="AluSx repeat: matches 1..303 of consensus"	
repeat_region	24518..24827	
	/note="AluSq repeat: matches 1..307 of consensus"	

```

repeat_region 26695..26785
/note="MIR repeat: matches 65..159 of consensus"
repeat_region 27638..27859
/note="MIR repeat: matches 13..261 of consensus"
repeat_region 28442..28509
/note="Tigger4(Zombi) repeat: matches 2663..2730 of consensus"
repeat_region 28510..28904
/note="MIR repeat: matches 115..509 of consensus"
repeat_region 29356..29650
/note="AluB repeat: matches 31..292 of consensus"
repeat_region 29793..29893
/note="match: GSS: Em.AQ469561"
repeat_region 32523..32677
/note="MIR repeat: matches 72..237 of consensus"
repeat_region 33683..33780
/note="L2 repeat: matches 2645..2747 of consensus"
repeat_region 34069..34561
/note="match: GSS: Em.B99319"
repeat_region 34328..34631
/note="AluSg repeat: matches 1..303 of consensus"
repeat_region 34765..34900
/note="MIR repeat: matches 31..165 of consensus"
repeat_region 35530..35560
/note="MIR repeat: matches 111..141 of consensus"
repeat_region 36192..36304
/note="MER81 repeat: matches 1..114 of consensus"
repeat_region 36313..36377
/note="MER5A repeat: matches 122..185 of consensus"
repeat_region 36473..36609
/note="MER5B repeat: matches 1..138 of consensus"
repeat_region 36849..37031
/note="LIME3 repeat: matches 5678..5855 of consensus"
repeat_region 37899..38062
/note="MER5A repeat: matches 26..198 of consensus"
repeat_region 38133..38708
/note="LMB5 repeat: matches 5584..6161 of consensus"
repeat_region 39180..39323
/note="MIR repeat: matches 93..250 of consensus"
repeat_region 40106..40418
/note="AluSx repeat: matches 1..312 of consensus"
repeat_region 40703..40785
/note="MIR repeat: matches 314..398 of consensus"
repeat_region 41356..41443
/note="MIR repeat: matches 46..142 of consensus"
repeat_region 43245..44512
/note="Charliela repeat: matches 169..1455 of consensus"
repeat_region 44513..44853
/note="MER7A repeat: matches 1..346 of consensus"
repeat_region 44854..45063
/note="Charliela repeat: matches 1..168 of consensus"
repeat_region 45167..45356
/note="L2 repeat: matches 1244..1442 of consensus"
repeat_region 45495..45911
/note="MIR repeat: matches 33..466 of consensus"
repeat_region 46045..46072
/note="L4 copies 2 mer gt 96% conserved"
repeat_region 46887..47198
/note="AluSx repeat: matches 1..312 of consensus"
repeat_region 48102..48173
/note="MIR repeat: matches 140..199 of consensus"
repeat_region 48174..48553
/note="THE1C repeat: matches 1..371 of consensus"
repeat_region 48554..48632
/note="MIR repeat: matches 57..140 of consensus"
repeat_region 49831..50014
/note="AluB repeat: matches 129..312 of consensus"
repeat_region 50076..50176
/note="MER69 repeat: matches 35..137 of consensus"
repeat_region 50349..50954
/note="MER21B repeat: matches 12..625 of consensus"

```

```

repeat_region 52068..52288
/note="MIR repeat: matches 2..239 of consensus"
repeat_region 52400..52455
/note="28 copies 2 mer ta 73% conserved"
repeat_region 52863..53272
/note="L2 repeat: matches 1696..2082 of consensus"
repeat_region 53273..53403
/note="FLAM C repeat: matches 1..132 of consensus"
repeat_region 53404..53726
Query Match 1.3% Score 21; DB 9; Length 128461;
Best Local Similarity 100.0%; Pred.No.16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1592 AAGCATCAGGCCAGACTGGGC 1612
DB 59698 AAGCATCAGGCCAGACTGGGC 59678

RESULT 85
AC022768 147124 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens clone RP11-338H7, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC022768
ACCESSION AC022768
VERSION AC022768.2 GI:7382466
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 147124)
Homo sapiens, clone RP11-338H7
Unpublished
2 (bases 1 to 147124)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Castle,A.,
Choepl,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howard,J.C., Johnson,K., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-PEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2000 this sequence version replaced gi:6922528.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5860
Center clone name: 338_H_7

```

* NOTE: This record contains 156 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
793: contig of 793 bp in length
794 893: gap of 100 bp
894 1668: contig of 775 bp in length
1669 1768: gap of 100 bp
1769 2609: contig of 841 bp in length
2610 2709: gap of 100 bp
2710 3509: contig of 800 bp in length
3510 3609: gap of 100 bp
3610 4428: contig of 819 bp in length
4429 4528: gap of 100 bp
4529 5307: contig of 779 bp in length
5308 5407: gap of 100 bp
5408 6197: contig of 790 bp in length
6198 6297: gap of 100 bp
6298 6986: contig of 689 bp in length
6987 7086: gap of 100 bp
7087 7838: contig of 752 bp in length
7839 7938: gap of 100 bp
7939 8707: contig of 769 bp in length
8708 8807: gap of 100 bp
8808 9603: contig of 796 bp in length
9604 9703: gap of 100 bp
9704 10510: contig of 807 bp in length
10511 10610: gap of 100 bp
10611 11388: contig of 778 bp in length
11389 11488: gap of 100 bp
11489 12295: contig of 807 bp in length
12296 12395: gap of 100 bp
12396 13230: contig of 835 bp in length
13231 13330: gap of 100 bp
13331 14176: contig of 846 bp in length
14177 14276: gap of 100 bp
14277 15045: contig of 769 bp in length
15046 15145: gap of 100 bp
15146 15917: contig of 772 bp in length
15918 16017: gap of 100 bp
16018 16777: contig of 760 bp in length
16778 16877: gap of 100 bp
16878 17673: contig of 796 bp in length
17674 17773: gap of 100 bp
17774 18524: contig of 751 bp in length
18525 18624: gap of 100 bp
18625 19423: contig of 799 bp in length
19424 19523: gap of 100 bp
19524 20326: contig of 803 bp in length
20327 20426: gap of 100 bp
20427 21202: contig of 776 bp in length
21203 21302: gap of 100 bp
21303 22069: contig of 767 bp in length
22070 22169: gap of 100 bp
22170 22971: contig of 802 bp in length
22972 23071: gap of 100 bp
23072 23876: contig of 805 bp in length
23877 23976: gap of 100 bp
23977 24774: contig of 798 bp in length
24775 24874: gap of 100 bp
24875 25650: contig of 776 bp in length
25651 25750: gap of 100 bp
25751 26542: contig of 792 bp in length
26543 26642: gap of 100 bp
26643 27433: contig of 791 bp in length
27434 27533: gap of 100 bp
27534 28325: contig of 792 bp in length
28326 28425: gap of 100 bp
28426 29235: contig of 810 bp in length
29236 29335: gap of 100 bp
29336 30134: contig of 799 bp in length

30135 30234: gap of 100 bp
30235 31044: contig of 810 bp in length
31045 31144: gap of 100 bp
31145 31923: contig of 779 bp in length
31924 32023: gap of 100 bp
32024 32804: contig of 781 bp in length
32805 32904: gap of 100 bp
32905 33648: contig of 744 bp in length
33649 33748: gap of 100 bp
33749 34550: contig of 802 bp in length
34551 34650: gap of 100 bp
34651 35398: contig of 748 bp in length
35399 35498: gap of 100 bp
35499 36304: contig of 806 bp in length
36305 36404: gap of 100 bp
36405 37202: contig of 798 bp in length
37203 37302: gap of 100 bp
37303 38139: contig of 837 bp in length
38140 38239: gap of 100 bp
38240 39045: contig of 806 bp in length
39046 39145: gap of 100 bp
39146 39931: contig of 786 bp in length
39932 40031: gap of 100 bp
40032 40830: contig of 799 bp in length
40831 40930: gap of 100 bp
40931 41701: contig of 771 bp in length
41702 41801: gap of 100 bp
41802 42585: contig of 784 bp in length
42586 42686: gap of 100 bp
42687 43490: contig of 805 bp in length
43491 43590: gap of 100 bp
43591 44360: contig of 770 bp in length
44361 44460: gap of 100 bp
44461 45263: contig of 803 bp in length
45264 45363: gap of 100 bp
45364 46147: contig of 783 bp in length
46148 46246: gap of 100 bp
46247 47065: contig of 819 bp in length
47066 47165: gap of 100 bp
47166 47956: contig of 791 bp in length
47957 48056: gap of 100 bp
48057 48833: contig of 777 bp in length
48834 48933: gap of 100 bp
48934 49716: contig of 783 bp in length
49717 49816: gap of 100 bp
49817 50577: contig of 761 bp in length
50578 50677: gap of 100 bp
50679 51461: contig of 784 bp in length
51462 51561: gap of 100 bp
51562 52351: contig of 790 bp in length
52352 52451: gap of 100 bp
52452 53218: contig of 767 bp in length
53219 53318: gap of 100 bp
53319 54114: contig of 796 bp in length
54115 54214: gap of 100 bp
54215 55025: contig of 811 bp in length
55026 55125: gap of 100 bp
55126 55943: contig of 818 bp in length
55944 56043: gap of 100 bp
56044 56880: contig of 837 bp in length
56881 56980: gap of 100 bp
56981 57775: contig of 795 bp in length
57776 57875: gap of 100 bp
57876 58675: contig of 800 bp in length
58676 58775: gap of 100 bp
58776 59557: contig of 782 bp in length
59558 59657: gap of 100 bp
59658 60437: contig of 780 bp in length
60438 60537: gap of 100 bp
60538 61326: contig of 789 bp in length
61327 61426: gap of 100 bp
61427 62225: contig of 799 bp in length
62226 62325: gap of 100 bp

* 62326 63098: contig of 773 bp in length
 * 63099 63198: gap of 100 bp

Query Match 1.3%; Score 21; DB 2; Length 147124;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 AGAAGCTGGCTGCTCTGTGCA 580

Db 130942 AGAAGCTGGCTGCTCTGTGCA 130962

RESULT 86

AC118107 157032 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-308P4, *** SEQUENCING IN PROGRESS
 DEFINITION ***; 2 unordered pieces.

AC118107 AC118107.6 GI:25009416

KEYWORDS HTG; HTGS PHASR1; HTGS DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 157032)
 Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyatebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganra, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nair, L., Nwankwelu, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. O., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished
 2 (bases 1 to 157032)

Worley, K. C.

Direct Submission

Submitted (13-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 157032)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23194896.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GVEH

Center clone name: CH230-308P4

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 152919 bases at least Q40

Consensus quality: 154343 bases at least Q30

Consensus quality: 155048 bases at least Q20

Estimated insert size: 160027; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTES

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 3824: contig of 3824 bp in length

* 3825 3924: gap of unknown length

* 3925 157032: contig of 153108 bp in length.

Location/Qualifiers

1. 157032

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clones="CH230-308P4"

1. 1929

/note="wgs end extension

clone_end:T7"

2226..3824

/note="wgs end extension

clone_end:T7"

2978..3872

/note="clone boundary

clone_end:T7"

FEATURES

source

misc_feature

clone_end:T7"

2226..3824

misc_feature

clone_end:T7"

2978..3872

misc_feature

clone_end:T7"

```

misc_feature      site:
end_sequence:B2279827"
3925..6208
/note="wgs contig"
misc_feature      47084..47813
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:B2279828"

ORIGIN
Query Match      1.3%; Score 21; DB 2; Length 157032;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 CCAACAATCCAAATCCAAAAA 52
|||||
Db 27539 CCAACAATCCAAATCCAAAAA 27559

RESULT 87
AC079118/c
LOCUS      158801 bp DNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens BAC clone RP11-260K18 from 4, complete sequence.
ACCESSION  AC079118
VERSION     AC079118.5 GI:15638800
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 158801)
AUTHORS   Sulston, J.E. and Waterston, R.
TITLE     Toward a complete human genome sequence
JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE   99063792
PUBMED    9847074
REFERENCE  2 (bases 1 to 158801)
AUTHORS   Cordes, M., Meyer, R. and Doebber, A.
TITLE     The sequence of Homo sapiens BAC clone RP11-260K18
JOURNAL   Unpublished
REFERENCE  3 (bases 1 to 158801)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (18-AUG-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  4 (bases 1 to 158801)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (18-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  5 (bases 1 to 158801)
AUTHORS   Waterston, R.
TITLE     Direct Submission
JOURNAL   Submitted (07-NOV-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Sep 18, 2001 this sequence version replaced gi:13876551.
            ----- Genome Center
            -----
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_NH0260K18
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-415L23; the clone sequenced to the right is AC036224. Actual start of this clone is at base position 1 of RP11-260K18; actual end is at base position 158801 of RP11-260K18.

FEATURES

Location/Qualifiers

```

1..158801
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-260K18"
/clone_lib="RP11-11"
11..300
/rpt_family="L2"
301..327
/rpt_family="(T)n"
609..930
/rpt_family="L2"
1014..1534
/rpt_family="L2"
2876..3009
/note="match to EST BG753882 (NID:G14064535)"
3326..3629
/rpt_family="ERV"
3717..3874
/rpt_family="MIR"
3948..3979
/note="match to EST BG753882 (NID:G14064535)"
5219..5313
/rpt_family="L1"
5314..5632
/rpt_family="Alu"
5633..6225
/rpt_family="L1"
5869..5927
/rpt_family="AT-rich"
6221..6240
/note="match to EST AW302415 (NID:G6712092) xr46f04.x1"
6228..6539
/rpt_family="Alu"
6554..6850
/rpt_family="Alu"
7037..7205
/rpt_family="L1"
7206..7271
/rpt_family="Mariner"
7272..7484

```

```

/rpt_family="L1"
7485. .7762
/rpt_family="Alu"
7763. .8132
/rpt_family="L1"
8310. .8602
/rpt_family="L1"
8708. .9000
/rpt_family="Alu"
8977. .9000
/rpt_family="(A)n"
9160. .9812
/rpt_family="L1"
9720. .9742
/rpt_family="AT_rich"
10167. .10383
/rpt_family="L1"
10386. .10469
/rpt_family="AT_rich"
10757. .10801
/rpt_family="AT_rich"
10770. .11083
/rpt_family="Alu"
11131. .11658
/rpt_family="L1"
12726. .13011
/rpt_family="Alu"
13011. .13401
/rpt_family="(TA)n"
13401. .13422
/rpt_family="(GA)n"
13702. .13760
/notes="match to EST BG483797 (NID:gl3415986)"
13771. .13810
/rpt_family="AT_rich"
13918. .13945
/rpt_family="AT_rich"
14250. .14511
/rpt_family="L1"
14484. .14505
/rpt_family="AT_rich"
14534. .14687
/rpt_family="L2"
16345. .16552
/rpt_family="L1"
16586. .16901
/rpt_family="Alu"
16705. .16730
/rpt_family="(A)n"
16878. .16900
/rpt_family="(TAAA)n"
17189. .17331
/rpt_family="MIR"
17360. .17414
/rpt_family="MIR"
17989. .18130
/rpt_family="Alu"
18796. .18874
/rpt_family="MIR"
19343. .19408
/notes="match to EST BG483797 (NID:gl3415986)"
19726. .19781
/rpt_family="L1"
19781. .20021
/rpt_family="L1"
20509. .20644
/rpt_family="MIR1_type"
20829. .20973
/rpt_family="MIR"
20917. .21159
/notes="match to EST BF870694 (NID:gl2260824)"
21112. .21317
/notes="match to EST BG483797 (NID:gl3415986)"

```

```

repeat_region 21872..21930
misc_feature   /rpt_family="L2"
                22012..22114

Query Match      1.3%; Score 21; DB 9; Length 158801;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1098 AGTAGAGTCCAGAGGATGTG 1118
      |||||
DB 132788 AGTAGAGTCCAGAGGATGTG 132768

RESULT 88
AC134703
LOCUS          Rattus norvegicus clone CH230-499J12, WORKING DRAFT SEQUENCE, 2
DEFINITION    unordered pieces.
ACCESSION     AC134703
VERSION       AC134703.2 GI:25138700
KEYWORDS      HTG; HTGS_PHRASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
               Rattus.
REFERENCE     1 (bases 1 to 159281)
AUTHORS       Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
               Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angulano,D.,
               Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
               Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
               Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
               Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
               Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
               Chacko,J., Chavez,B., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
               Cleveland,C., Cockrell,R., Cox,C., Coyte,M., Cree,A., D'Souza,L.,
               Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
               Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
               Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
               Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
               Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
               Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
               Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
               Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
               Harvey,Y., Ravlak,P., Hawes,A., Henderson,N., Hernandez,J.,
               Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
               Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
               Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
               Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Liu,J.,
               Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
               Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
               Lorensuhewa,L., Loulsegid,H., Lozado,R.J., Lu,X., Ma,J.,
               Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
               Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
               Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
               Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
               Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
               Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
               Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
               Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
               Plummer,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
               Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
               Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
               Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Ruiz,S.J.,
               Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
               Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Sma's,D.,
               Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
               Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
               Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
               Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
               Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,E.,
               Williams,G., Willson,R., Wleczyk,R., Woodden,H., Worley,K.,
               Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

```

Yu, F., Zhang, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R.A., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished

2 (bases 1 to 159281)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (30-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 159281)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 20, 2002 this sequence version replaced gi:23346380.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KCV4

Center clone name: CH230-499J12

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 155117 bases at least Q40

Consensus quality: 156153 bases at least Q30

Consensus quality: 156834 bases at least Q20

Estimated insert size: 159436; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 154891: contig of 154891 bp in length

* 154892 154991: gap of unknown length

* 154992 159281: contig of 4290 bp in length.

Location/Qualifiers

1. 159281

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-499J12"

31460..31625

/note="clone boundary

clone_end:Sp6

site:

end_sequence:B271251"

ORIGIN

Query Match 1.3%; Score 21; DB 2; Length 159281;

Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1166 TATACACAGAGAGTTTGAGG 1186

Db 84495 TATACACAGAGAGTTTGAGG 84515

RESULT 89

AC108791

LOCUS

DEFINITION

Mus musculus clone RP24-292D18, WORKING DRAFT SEQUENCE, 3 unordered

pieces.

ACCESSION

AC108791.4 GI:29126416

KEYWORDS

HTG; HTGS; PHASE1; HTGS_DRAFT.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 161728)

2 (bases 1 to 161728)

Unpublished

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,

Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,

Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,

McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,

Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,

Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,

Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 161728)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,

Boguslavsky, L., Bouckhgalter, B., Camarata, J., Chang, J., Choepel, Y.,

Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,

Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,

Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,

Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,

Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,

Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,

Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,

Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,

Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,

Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,

Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (20-MAR-2003) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 20, 2003 this sequence version replaced gi:28273548.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L20062
 Center clone name: 292_D_18
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 161128 bases at least Q40
 Consensus quality: 161507 bases at least Q30
 Insert size: 157000; agarose-fp
 Insert size: 161528; sum-of-contigs
 Quality coverage: 9.8 in Q20 bases; agarose-fp
 Quality coverage: 9.5 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 73166: contig of 73166 bp in length
* 73167 73265: gap of 100 bp
* 73267 142750: contig of 69484 bp in length
* 142751 142850: gap of 100 bp
* 142851 161728: contig of 18878 bp in length.
```

FEATURES

```

source
1. 161728
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /clone="RP24-292D18"
  /clone_lib="RPCI-24 Male Mouse BAC"
misc_feature
1. 73166
  /note="assembly_fragment"
  /clone_end:SP6
  vector_side:left"
misc_feature
73267..142750
  /note="assembly_fragment"
misc_feature
142851..161728
  /note="assembly_fragment"
  /clone_end:T7
  vector_side:right"
```

ORIGIN

```

Query Match      1.3%  Score 21;  DB 2;  Length 161728;
Best Local Similarity 100.0%;  Pred. No. 16;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
QY 1217 AGGTATTTCACCATTCACGC 1237
```

```
Db 125697 AGGTATTTCACCATTCACGC 125717
```

RESULT 90

```

AC110054/c
LOCUS AC110054 164857 bp DNA linear PRI 01-MAY-2002
DEFINITION Homo sapiens chromosome 15, clone CTD-2116N17, complete sequence.
ACCESSION AC110054
VERSION AC110054.4 GI:20377046
KEYWORDS HTG.
```

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 164857)

Biren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 15, clone CTD-2116N17

Unpublished

2 (bases 1 to 164857)

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,

Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,

Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,

McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,

Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,

Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,

Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S.,

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (09-FEB-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 164857)

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,

Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,

Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,

McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,

Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,

Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,

Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S.,

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 164857)

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,

Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,

Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Fato, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,

MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,

TITLE JOURNAL REFERENCE AUTHORS

Homo sapiens chromosome 15, clone RP11-537K8
Unpublished
2 (bases 1 to 169942)
Birren,B., Linton,L., Nubaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kanat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Marguis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoli,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talanas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169942)
Birren,B., Linton,L., Nubaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kanat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Marguis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoli,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talanas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 20, 2002 this sequence version replaced gi:21431214.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL REFERENCE AUTHORS

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27369
Center clone name: 537_K_8
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 163325 bases at least Q40
Consensus quality: 166531 bases at least Q30
Consensus quality: 167587 bases at least Q20

TITLE JOURNAL COMMENT

Insert size: 172000; agarose-fp
Insert size: 168542; sum-of-contigs
Quality coverage: 6.7 in Q20 bases; agarose-fp
Quality coverage: 6.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1656: contig of 1656 bp in length
* 1657 1756: gap of 100 bp
* 1757 2477: contig of 721 bp in length
* 2478 2577: gap of 100 bp
* 2578 3879: contig of 1302 bp in length
* 3880 3979: gap of 100 bp
* 3980 5480: contig of 1501 bp in length
* 5481 5580: gap of 100 bp
* 5581 7330: contig of 1750 bp in length
* 7331 7430: gap of 100 bp
* 7431 42565: contig of 35135 bp in length
* 42566 45761: contig of 3096 bp in length
* 45762 45861: gap of 100 bp
* 45862 50498: contig of 4637 bp in length
* 50499 50598: gap of 100 bp
* 50599 63026: contig of 12428 bp in length
* 63027 81162: contig of 18036 bp in length
* 81163 81262: gap of 100 bp
* 81263 95306: contig of 14044 bp in length
* 95307 95406: gap of 100 bp
* 95407 114859: contig of 19453 bp in length
* 114860 141337: contig of 26378 bp in length
* 141338 141437: gap of 100 bp
* 141438 167338: contig of 25901 bp in length
* 167339 167438: gap of 100 bp
* 167439 169942: contig of 2504 bp in length.
Location/Qualifiers
1..169942
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-537K8"
/clone_lib="RPC1-11 Human Male BAC"
1..1656
/note="assembly_fragment"
clone end:SP6
vector side:left
1757..2477
/note="assembly_fragment"
2578..3879
/note="assembly_fragment"
3980..5480
/note="assembly_fragment"
5581..7330
/note="assembly_fragment"
7431..42565
/note="assembly_fragment"
42566..45761
/note="assembly_fragment"
45862..50498
/note="assembly_fragment"
50599..63026
/note="assembly_fragment"
63027..81162
/note="assembly_fragment"

```

misc_feature 81263..95306 /note="assembly_fragment"
misc_feature 95407..114859 /note="assembly_fragment"
misc_feature 114960..141337 /note="assembly_fragment"
misc_feature 141438..167338 /note="assembly_fragment"
misc_feature 167439..169942 /note="assembly_fragment"
               clone_end:T7
               vector_side:right"

ORIGIN
Query Match      1.3%; Score 21; DB 2; Length 169942;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCGCCGACGACG 147
Db 93839 GCAGAGGTCGCCGACGACG 93819

RESULT 92
AC124792/c 173043 bp DNA linear HTG 17-JUN-2002
LOCUS Homo sapiens chromosome 15 clone RP11-537K8 map 15, WORKING DRAFT
DEFINITION SEQUENCE, 14 ordered pieces.
ACCESSION AC124792
VERSION AC124792.1 GI:21431213
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 173043)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 15, clone RP11-537K8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 173043)
          Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
          Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
          Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
          Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A.,
          Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
          Fark, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,
          Galagan, J., Gardyna, S., Ginde, S., Cord, S., Goyette, M., Graham, L.,
          Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
          Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRoque, K.,
          Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
          Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,
          Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
          Menais, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
          Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
          O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
          Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
          Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
          Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
          Stojanovic, N., Strauss, N., Subramanian, A., Talanas, J., Tesfaye, S.,
          Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Tesfaye, S.,
          Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
          Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
          Direct Submission
          Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          All repeats were identified using RepeatMasker:
          Smit, A.P.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RW/RepeatMasker.html
          ----- Genome Center
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu

```

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L27368

Center clone name: 537 K.8

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 168964 bases at least Q40

Consensus quality: 170311 bases at least Q30

Consensus quality: 170917 bases at least Q20

Insert size: 172000; agarose-fp

Insert size: 171743; sum-of-contigs

Quality coverage: 6.9 in Q20 bases; agarose-fp

Quality coverage: 6.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 14 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 15693: contig of 15693 bp in length

* 15694 15793: gap of 100 bp

* 15794 17371: contig of 1578 bp in length

* 17372 17471: gap of 100 bp

* 17472 19298: contig of 1827 bp in length

* 19299 19398: gap of 100 bp

* 19399 20994: contig of 1596 bp in length

* 20995 21094: gap of 100 bp

* 21095 21156: contig of 2062 bp in length

* 21157 22256: gap of 100 bp

* 22257 26786: contig of 3530 bp in length

* 26787 26886: gap of 100 bp

* 26887 33049: contig of 6163 bp in length

* 33050 33149: gap of 100 bp

* 33150 46188: contig of 13039 bp in length

* 46189 46288: gap of 100 bp

* 46289 60957: contig of 14669 bp in length

* 60958 61057: gap of 100 bp

* 61058 79888: contig of 18831 bp in length

* 79889 79988: gap of 100 bp

* 79989 95915: contig of 15927 bp in length

* 95916 96015: gap of 100 bp

* 96016 118578: contig of 22563 bp in length

* 118579 118678: gap of 100 bp

* 118679 144470: contig of 25792 bp in length

* 144471 144570: gap of 100 bp

* 144571 173043: contig of 28473 bp in length.

Location/Qualifiers

1. 173043

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-537K8"

/clone_lib="RPC1-11 Human Male BAC"

1..15693

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

15794..17371

/note="assembly_fragment"

17472..19298

/note="assembly_fragment"

19399..20994

/note="assembly_fragment"

21095..23156

/note="assembly_fragment"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

FEATURES

source

```

misc_feature 23257..26786
              /note="assembly_fragment"
misc_feature 26887..33049
              /note="assembly_fragment"
misc_feature 33150..46188
              /note="assembly_fragment"
misc_feature 46289..60957
              /note="assembly_fragment"
misc_feature 61058..79888
              /note="assembly_fragment"
misc_feature 79989..95915
              /note="assembly_fragment"
misc_feature 96016..118578
              /note="assembly_fragment"
misc_feature 118679..144470
              /note="assembly_fragment"
misc_feature 144571..173043
              /note="assembly_fragment"
              clone_end:T7
              vector_side:right"

ORIGIN
Query Match      1.3%; Score 21; DB 2; Length 173043;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGCAGC 147
    |||||
Db 14201 GCAGAGGTCCTCCGCGCAGC 14181

RESULT 93
AC100840/c
LOCUS AC100840 174287 bp DNA linear PRI 07-MAR-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-537K8, complete sequence.
ACCESSION AC100840
VERSION AC100840.2 GI:192225049
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 174287)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Lamaze,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 7, 2002 this sequence version replaced gi:17048210.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Center project name: L21571
Center clone name: 537_K_8
-----
Location/Qualifiers
1..174287
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-537K8"
/clone_lib="RPCI-11 Human Male BAC"
754..794
/rpt_family="CT-rich"
repeat_region
3180..3482
/rpt_family="AluSx"
repeat_region
7561..7859
/rpt_family="AluSx"
repeat_region
9724..9777
/rpt_family="T-rich"
repeat_region
10503..10561
/rpt_family="TG)n"
repeat_region
10825..10847
/rpt_family="AT-rich"
repeat_region
11581..11612
/rpt_family="AT-rich"
repeat_region
11613..11905
/rpt_family="AluSx"
repeat_region
complement(12336..12615)
/rpt_family="AluDb"
repeat_region
13968..13996
/rpt_family="GC-rich"
repeat_region
15225..15352
/rpt_family="AluJo"
repeat_region
complement(15353..15649)
/rpt_family="AluSq"
repeat_region
15650..15831
/rpt_family="AluJo"

```

```

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Lamaze,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (07-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 7, 2002 this sequence version replaced gi:17048210.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Center project name: L21571
Center clone name: 537_K_8
-----
Location/Qualifiers
1..174287
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-537K8"
/clone_lib="RPCI-11 Human Male BAC"
754..794
/rpt_family="CT-rich"
repeat_region
3180..3482
/rpt_family="AluSx"
repeat_region
7561..7859
/rpt_family="AluSx"
repeat_region
9724..9777
/rpt_family="T-rich"
repeat_region
10503..10561
/rpt_family="TG)n"
repeat_region
10825..10847
/rpt_family="AT-rich"
repeat_region
11581..11612
/rpt_family="AT-rich"
repeat_region
11613..11905
/rpt_family="AluSx"
repeat_region
complement(12336..12615)
/rpt_family="AluDb"
repeat_region
13968..13996
/rpt_family="GC-rich"
repeat_region
15225..15352
/rpt_family="AluJo"
repeat_region
complement(15353..15649)
/rpt_family="AluSq"
repeat_region
15650..15831
/rpt_family="AluJo"

```

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

DEFINITION Homo sapiens chromosome 15 clone RP11-537K8, WORKING DRAFT
 SEQUENCE, 11 unordered pieces.
 ACCESSION AC017098
 VERSION AC017098.2 GI:7024110
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 175258)
 AUTHORS Waterston,R.H.
 JOURNAL Direct Submission
 TITLE Submitted (09-DEC-1999) Genome Sequencing Center, Washington
 UNIVERSITY School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Feb 23, 2000 this sequence version replaced gi:5554044.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H_NH0537K08
 ----- Summary Statistics -----
 Sequencing vector: M13; 84%
 Chemistry: Dye-primer ET; 84% of reads
 Chemistry: Dye-terminator Big Dye; 16% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 168842 bases at least Q40
 Consensus quality: 170783 bases at least Q30
 Consensus quality: 171770 bases at least Q20
 Insert size: 178000; agarose-fp
 Quality coverage: 5.78 in Q20 bases; agarose-fp
 Quality coverage: 5.90 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 3574: contig of 3574 bp in length
 * 3575 3674: gap of unknown length
 * 3675 3675: contig of 6218 bp in length
 * 9892 9892: gap of unknown length
 * 9932 9932: contig of 10396 bp in length
 * 20388 20388: gap of unknown length
 * 20489 33533: contig of 13045 bp in length
 * 33533 33533: gap of unknown length
 * 33634 33634: contig of 13496 bp in length
 * 47130 47229: gap of unknown length
 * 47229 62600: contig of 15371 bp in length
 * 62600 62700: gap of unknown length
 * 62701 79334: contig of 16634 bp in length
 * 79334 79335: gap of unknown length
 * 79335 79879: contig of 18445 bp in length
 * 79879 97979: gap of unknown length
 * 97979 97980: contig of 19588 bp in length
 * 97980 117567: gap of unknown length
 * 117567 117667: gap of unknown length
 * 117668 144817: contig of 27150 bp in length
 * 144818 144917: gap of unknown length
 * 144918 175258: contig of 30341 bp in length.
 Location/Qualifiers
 1. 175258
 /organism="Homo sapiens"

FEATURES
 source

/mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /clones="RP11-537K8"
 1. 3574
 /note="assembly_name:Contig6"
 3675. 9892
 /note="assembly_name:Contig7"
 9993. 20388
 /note="assembly_name:Contig8"
 20489. 33533
 /note="assembly_name:Contig9"
 33634. 47129
 /note="assembly_name:Contig10
 clone_end:SP6
 vector_side:right
 47230. 62600
 /note="assembly_name:Contig11"
 62701. 79334
 /note="assembly_name:Contig12"
 79435. 97879
 /note="assembly_name:Contig13"
 97980. 117567
 /note="assembly_name:Contig14"
 117668. 144817
 /note="assembly_name:Contig15"
 144918. 175258
 /note="assembly_name:Contig16"

ORIGIN

Query Match 1.3%; Score 21; DB 2; Length 175258;
 Best Local Similarity 100.0%; Pred.No.16;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 GCAGAGGTCCTCCGCGCAGC 147
 |||||
 DB 19382 GCAGAGGTCCTCCGCGCAGC 19402

RESULT 96
 AC067846/c

LOCUS AC067846 178344 bp DNA linear HTG 24-OCT-2001
 DEFINITION Homo sapiens chromosome 15 clone RP11-336A3 map 15, WORKING DRAFT
 SEQUENCE, 16 unordered pieces.

ACCESSION AC067846

VERSION AC067846.3 GI:16356918

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 178344)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 15, clone RP11-336A3

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 178344)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavskiy,L., Bouckhagalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domina,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlungu,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (27-Apr-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 24, 2001 this sequence version replaced gi:12061542.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9366

Center clone name: 336_A_3

----- Summary Statistics

Sequencing vector: M13; M77815; 3% of reads
 Sequencing vector: Plasmid; n/a; 97% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 173722 bases at least Q40
 Consensus quality: 175568 bases at least Q30
 Consensus quality: 176332 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 176844; sum-of-contigs
 Quality coverage: 8.5 in Q20 bases; agarose-fp
 Quality coverage: 8.2 in Q20 bases

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 6756: contig of 6756 bp in length
 * 6757 6856: gap of 100 bp
 * 6857 7994: contig of 1138 bp in length
 * 7995 8094: gap of 100 bp
 * 8095 9256: contig of 1162 bp in length
 * 9257 9356: gap of 100 bp
 * 9357 10821: contig of 1495 bp in length
 * 10822 10951: gap of 100 bp
 * 10952 51439: contig of 40488 bp in length
 * 51440 51539: gap of 100 bp
 * 51540 52253: contig of 714 bp in length
 * 52254 52353: gap of 100 bp
 * 52354 53099: contig of 746 bp in length
 * 53100 53199: gap of 100 bp
 * 53200 59086: contig of 5887 bp in length
 * 59087 59186: gap of 100 bp
 * 59187 59932: contig of 746 bp in length
 * 59933 60032: gap of 100 bp
 * 60033 66468: contig of 6436 bp in length
 * 66469 66568: gap of 100 bp
 * 66569 77105: contig of 10537 bp in length
 * 77106 77205: gap of 100 bp
 * 77206 96872: contig of 19667 bp in length
 * 96873 120255: contig of 23283 bp in length
 * 120256 120355: gap of 100 bp
 * 120356 144957: contig of 24602 bp in length
 * 144958 145058: gap of 100 bp
 * 145059 177826: contig of 32769 bp in length
 * 177827 177927: gap of 100 bp
 * 177928 178344: contig of 418 bp in length.

----- Location/Qualifiers

1..178344

/organism="Homo sapiens"

/mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-336A3"
 /clone_lib="RPC1-11 Human Male BAC"
 1..6756
 /note="assembly_fragment
 clone end:SP6
 vector side:left"
 misc_feature
 6857..7994
 /note="assembly_fragment"
 misc_feature
 8095..9256
 /note="assembly_fragment"
 misc_feature
 9357..10851
 /note="assembly_fragment"
 misc_feature
 10952..51439
 /note="assembly_fragment"
 misc_feature
 51540..52253
 /note="assembly_fragment"
 misc_feature
 52354..53099
 /note="assembly_fragment"
 misc_feature
 53200..59086
 /note="assembly_fragment"
 misc_feature
 59187..59932
 /note="assembly_fragment"
 misc_feature
 60033..66468
 /note="assembly_fragment"
 misc_feature
 66569..77105
 /note="assembly_fragment"
 misc_feature
 77206..96872
 /note="assembly_fragment"
 misc_feature
 96973..120255
 /note="assembly_fragment"
 misc_feature
 120356..144957
 /note="assembly_fragment"
 misc_feature
 145058..177826
 /note="assembly_fragment"
 misc_feature
 177927..178344
 /note="assembly_fragment
 clone end:T7
 vector side:right"

ORIGIN

Query Match 1.3%; Score 21; DB 2; Length 178344;

Best Local Similarity 100.0%; Pred.No.16;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

Db 118739 GCAGAGGTCCCGCAGCAGC 118719

RESULT 97

AC024631

LOCUS

AC024631 178611 bp DNA linear HTG 27-MAR-2003

DEFINITION Homo sapiens chromosome 15 clone RP11-801F6 map 15, 3 unordered

ACCESSION

AC024631

VERSION

HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 178611)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 15, clone RP11-801F6

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 178611)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,

Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1. 11216: contig of 11216 bp in length
 * 11217 11316: gap of unknown length
 * 11317 22412: contig of 11096 bp in length
 * 22413 22512: gap of unknown length
 * 22513 41602: contig of 19090 bp in length
 * 41603 41702: gap of unknown length
 * 41703 64104: contig of 22302 bp in length
 * 64005 64104: gap of unknown length
 * 64105 90107: contig of 26003 bp in length
 * 90108 90207: gap of unknown length
 * 90208 183102: contig of 92895 bp in length.

FEATURES

source

1. 183102
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="RP44-182H14"
 /clone_lib="RP44"
 1. 11216
 /note="assembly_fragment"
 11317..22412
 /note="assembly_fragment"
 22513..41602
 /note="assembly_fragment"
 41703..64004
 /note="assembly_fragment"
 64105..90107
 /note="assembly_fragment"
 90208..183102
 /note="assembly_fragment"

ORIGIN

Query Match 1.3%; Score 21; DB 2; Length 183102;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 GGAGCCAGGAGCGGCCGAGC 105

Db 129213 GGAGCCAGGAGCGGCCGAGC 129193

RESULT 99

AC146969

LOCUS AC146969 211601 bp DNA linear HTG 30-OCT-2003

DEFINITION Sus scrofa clone RP44-20B15, WORKING DRAFT SEQUENCE, 6 unordered pieces.

ACCESSION AC146969

VERSION AC146969.1 GI:38044168

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.

SOURCE Sus scrofa (pig)

ORGANISM

Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Sutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 211601)

REFERENCE
 AUTHORS Antonellis A., Ayele K., Benjamin B., Blakesley R.W.,
 Bouffard G.G., Brinkley C., Brooks S., Chu G., Coleman B.,
 Coleman H., Engle J., Granite S., Guan X., Gupta J., Haghighi, P.,
 Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R.,
 Karlins, E., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-O.,
 Legaspi, R., Maduro, Q.B., Maduro, V.B., Margulies, E.H., Mastello, C.,
 Maskeri, B., McDowell, J., Mullikin, J.C., Paguirigan, C., Pearson, R.,
 Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K.,
 Schueler, M.G., Shah, K., Sison, C., Stantipop, S., Thomas, J.W.,
 Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and
 Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished

REFERENCE

2 (bases 1 to 211601)

Green, E.D.

Direct Submission

Submitted (30-OCT-2003) NIH Intramural Sequencing Center, 8717

Grovemont Circle, Gaithersburg, MD 20877, USA

COMMENT

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc.zoe@nih.gov

----- Project Information

Center project name: dyd

Center clone name: 020B15

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 208458 bases at least Q40

Consensus quality: 209373 bases at least Q30

Insert size: 208000; agarose-fp

Insert size: 211101; sum-of-contigs

Quality coverage: 12.36x in Q20 bases; agarose-fp

Quality coverage: 12.18x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2037: contig of 2037 bp in length
 * 2038 2137: gap of unknown length
 * 2138 16535: contig of 14398 bp in length
 * 16536 16635: gap of unknown length
 * 16636 35667: contig of 18931 bp in length
 * 35667 35668: gap of unknown length
 * 35669 61793: contig of 26127 bp in length
 * 61794 61894: gap of unknown length
 * 61894 120454: contig of 58561 bp in length
 * 120455 120554: gap of unknown length
 * 120555 211601: contig of 91047 bp in length.

FEATURES

source

1. 211601
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="RP44-20B15"
 /clone_lib="RP44"
 1. 2037
 /note="assembly_fragment"
 2138..16535
 /note="assembly_fragment"
 16636..35666
 /note="assembly_fragment"
 35667..61793
 /note="assembly_fragment"
 61894..120454
 /note="assembly_fragment"
 120555..211601
 /note="assembly_fragment"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 1.3%; Score 21; DB 2; Length 211601;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 GGAGCCAGGAGCGGCCGAGC 105

Db 115852 GGAGCCAGGAGCGGCCGAGC 115872

RESULT 100

AC106432/c

LOCUS

DEFINITION

AC106432

AC106432.5

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 212026)

REFERENCE

AUTHORS

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Amin, A., Angiano, D.,

Allen, C., Allen, H., Albrooks, S., Ayodeji, M., Baca, E., Baden, H.,

Ayalabechi, V., Ayagi, A., Ayodeji, M., Barnstead, M., Barnstead, F.,

Baldwin, D., Bandaranaike, D., Barber, M., Blyth, P., Brown, M.,

Biswal, K., Blair, J., Blankenburg, K., Burrell, K., Calderon, E.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davia, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,

Gunnaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hui, S., Hume, J., Idebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensueta, L., Louis, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nair, L.,

Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,

Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,

Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,

Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,

Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

Williams, G., Willson, R., Wlezyk, R., Woodden, H., Worley, K.,

Wright, D., Wright, K., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,

Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 212026)

REFERENCE

AUTHORS

Worley, K. C.

Direct Submission

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 212026)

AUTHORS

TITLE

JOURNAL

COMMENT

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24819132.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center of Medicine

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKXB

Center clone name: CH230-188113

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 203313 bases at least Q40

Consensus quality: 205311 bases at least Q30

Consensus quality: 206730 bases at least Q20

Estimated insert size: 212502; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 117657: contig of 117657 bp in length

* 117658 117757: gap of unknown length

* 117758 181359: contig of 63602 bp in length

* 181360 181459: gap of unknown length

* 181460 210912: contig of 29453 bp in length

* 210913 211012: gap of unknown length

* 211013 212026: contig of 1014 bp in length.

* Location/Qualifiers

1 212026

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clones="CH230-188113"

complement(539..1284)

/note="clone boundary

clone end:T7

site:EcoRI

end_sequence:BH328753"

complement(116484..117144)

/note="clone boundary

clone end:Sp6

site:EcoRI

end_sequence:BH328754"

ORIGIN

Query Match

Best Local Similarity 1.3%; Score 21; DB 2; Length 212026;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TITLE

JOURNAL

REFERENCE

AUTHORS

Worley, K. C.

Direct Submission

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 212026)

Search completed: June 8, 2004, 20:31:59
Job time : 6241.26 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 12:36:00 ; Search time 3992.48 Seconds
(without alignments)
12251.597 Million cell updates/sec

Title: US-10-023-523-46

Perfect score: 1638

Sequence: 1 atgaagaacacagacaaaa.....agccacactccgcagggcc 1638

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

EST:

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hcc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_nam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	623	38.0	759	14	CF995807
2	535	32.7	606	12	B1006817
3	485	29.6	570	10	BF829485
4	467	28.5	799	13	BU849604

5	457	27.9	817	12	BG761234
6	447	27.3	835	14	CD655044
7	419	25.6	932	12	AGENCOURT
8	379	23.1	890	13	AGENCOURT
9	375	22.9	970	13	AGENCOURT
10	357	21.8	1085	12	AGENCOURT
11	355	21.7	1318	12	AGENCOURT
12	301	18.4	1195	13	AGENCOURT
13	267	16.3	1039	12	AGENCOURT
14	241	14.7	965	10	AGENCOURT
15	231	14.1	281	12	AGENCOURT
16	206	12.6	338	10	AGENCOURT
17	178	10.9	336	10	AGENCOURT
18	145	8.9	525	9	AGENCOURT
19	138	8.4	297	13	AGENCOURT
20	133	8.1	268	13	AGENCOURT
21	130	7.9	605	14	AGENCOURT
22	116	7.1	443	9	AGENCOURT
23	112	6.8	388	9	AGENCOURT
24	109	6.7	280	10	AGENCOURT
25	100	6.1	124	10	AGENCOURT
26	97	5.9	414	9	AGENCOURT
27	90	5.5	850	28	AGENCOURT
28	81	4.9	254	9	AGENCOURT
29	79	4.8	223	10	AGENCOURT
30	65	4.0	273	10	AGENCOURT
31	63	3.8	429	10	AGENCOURT
32	61	3.7	152	13	AGENCOURT
33	53	3.2	628	14	AGENCOURT
34	51	3.1	413	12	AGENCOURT
35	50	3.1	414	14	AGENCOURT
36	50	3.1	427	9	AGENCOURT
37	50	3.1	427	12	AGENCOURT
38	50	3.1	428	9	AGENCOURT
39	50	3.1	472	10	AGENCOURT
40	47	2.9	527	14	AGENCOURT
41	47	2.9	539	13	AGENCOURT
42	47	2.9	551	14	AGENCOURT
43	47	2.9	580	10	AGENCOURT
44	47	2.9	620	10	AGENCOURT
45	47	2.9	647	14	AGENCOURT
46	47	2.9	653	13	AGENCOURT
47	47	2.9	699	13	AGENCOURT
48	47	2.9	710	14	AGENCOURT
49	47	2.9	716	14	AGENCOURT
50	47	2.9	725	13	AGENCOURT
51	47	2.9	727	13	AGENCOURT
52	47	2.9	735	13	AGENCOURT
53	47	2.9	744	13	AGENCOURT
54	47	2.9	750	14	AGENCOURT
55	47	2.9	770	14	AGENCOURT
56	47	2.9	775	14	AGENCOURT
57	47	2.9	779	14	AGENCOURT
58	47	2.9	798	14	AGENCOURT
59	47	2.9	806	14	AGENCOURT
60	47	2.9	819	10	AGENCOURT
61	47	2.9	900	12	AGENCOURT
62	46	2.8	754	14	AGENCOURT
63	46	2.8	778	14	AGENCOURT
64	45	2.7	942	13	AGENCOURT
65	41	2.5	284	10	AGENCOURT
66	41	2.5	654	13	AGENCOURT
67	41	2.5	723	13	AGENCOURT
68	41	2.5	857	14	AGENCOURT
69	41	2.5	861	12	AGENCOURT
70	41	2.5	954	13	AGENCOURT
71	40	2.4	210	10	AGENCOURT
72	40	2.4	414	9	AGENCOURT
73	39	2.4	727	14	AGENCOURT
74	38	2.3	409	10	AGENCOURT
75	38	2.3	608	14	AGENCOURT
76	38	2.3	675	14	AGENCOURT
77	38	2.3	821	12	AGENCOURT

CF995807

B1006817

BF829485

BU849604

AGENCOURT

78 466 10 B5601883 266978 MA
79 497 14 C167914 B0791G03-
80 725 13 B0613265 UI-M-EWO-
81 725 13 B0613265 UI-M-EWO-
82 725 13 B0613265 UI-M-EWO-
83 725 13 B0613265 UI-M-EWO-
84 725 13 B0613265 UI-M-EWO-
85 725 13 B0613265 UI-M-EWO-
86 725 13 B0613265 UI-M-EWO-
87 725 13 B0613265 UI-M-EWO-
88 725 13 B0613265 UI-M-EWO-
89 725 13 B0613265 UI-M-EWO-
90 725 13 B0613265 UI-M-EWO-
91 725 13 B0613265 UI-M-EWO-
92 725 13 B0613265 UI-M-EWO-
93 725 13 B0613265 UI-M-EWO-
94 725 13 B0613265 UI-M-EWO-
95 725 13 B0613265 UI-M-EWO-
96 725 13 B0613265 UI-M-EWO-
97 725 13 B0613265 UI-M-EWO-
98 725 13 B0613265 UI-M-EWO-
99 725 13 B0613265 UI-M-EWO-
100 725 13 B0613265 UI-M-EWO-

ALIGNMENTS

RESULT 1
CF995807 759 bp mRNA linear EST 25-NOV-2003
LOCUS AGENCOURT 16109164 NIH MGC 221 Homo sapiens cDNA clone
DEFINITION IMAGE:30708497 5', mRNA sequence.

ACCESSION CF995807
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 759)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabs@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM1074 row: m column: 18
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

SOURCE

1..759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708497"
/lab_host="PH10B Tona"
/clone_lib="NIH MGC 221"
Note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. Average insert size
4-5kb. Adaptors 5' (AATCGGCAGAGG)3' and 5' (CTCTGCGG)3' 3'. Linker sequence - GCGGCGCTGAGGCC T18.
Sequencing primers 3' end: 13 promoter primer 5'd
(ATTAACCTCTCAATAAGGA)3', 5' End: T7 promoter primer 5'd
(TAATACGACTACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

ORIGIN

Query Match 38.0%; Score 623; DB 14; Length 759;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 609 CCGAATTCACAGAGCAGATGAAGTCTCTACAGAAAGAGCAGCAGCTGTGCAAGA 668
DB 8 CCGAATTCACAGAGCAGATGAAGTCTCTACAGAAAGAGCAGCAGCTGTGCAAGA 67
QY 669 GAAGGACCACTCGCGGTGAGCAGCAGCAAGGCGCTCTGCGCCGAGCAAGCTTGAGAG 728
DB 68 GAAGGACCACTCGCGGTGAGCAGCAGCAAGGCGCTCTGCGCCGAGCAAGCTTGAGAG 127
QY 729 CCTATGCGGTGAGTGCAGCGGCACACCGCTCCCTCAAGAGAGAGGTGTGAGCGGGC 788
DB 128 CCTATGCGGTGAGTGCAGCGGCACACCGCTCCCTCAAGAGAGAGGTGTGAGCGGGC 187
QY 789 CCGGAGGAGGAGAGGAGGAGGAGGAGTGCCTGCGACTTCCAGGTGACACTGAATGA 848
DB 188 CCGGAGGAGGAGAGGAGGAGGAGGAGTGCCTGCGACTTCCAGGTGACACTGAATGA 247
QY 849 CATTGAGCTCAGATGGAAGCAGCAGCAATGAGCGCAACTCCAGCTGCGCCAGAGACAT 908
DB 248 CATTGAGCTCAGATGGAAGCAGCAGCAATGAGCGCAACTCCAGCTGCGCCAGAGACAT 307
QY 909 GGAGCTGCGTGCAGAGCTCAAGAGGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATAT 968
DB 308 GGAGCTGCGTGCAGAGCTCAAGAGGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATAT 367
QY 969 CGACAAAGTCTTCAACACAGAGCAGCTTACAGAGCTGCTGATGCCAGCTCCAGCA 1028
DB 368 CGACAAAGTCTTCAACACAGAGCAGCTTACAGAGCTGCTGATGCCAGCTCCAGCA 427
QY 1029 GCGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1088
DB 428 GCGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 487
QY 1089 GAAAGAGGAGTATGAGTCCAGAGGAGTGTGAGCTGATGAAGCAGCAGAGACCCACCT 1148
DB 488 GAAAGAGGAGTATGAGTCCAGAGGAGTGTGAGCTGATGAAGCAGCAGAGACCCACCT 547
QY 1149 GAAGCAGCAGTTCCTTACACAGAGAGTGTGAGGAGTTCAGAGCAGCAGCTTTCCAA 1208
DB 548 GAAGCAGCAGTTCCTTACACAGAGAGTGTGAGGAGTTCAGAGCAGCAGCTTTCCAA 607
QY 1209 AAGCAGCAGTTCCTTACACAGAT 1231
DB 608 AAGCAGCAGTTCCTTACACAGAT 630

RESULT 2

BI006817

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
MEDLINE
PUBMED

20202663

10737800

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-RT0067-150101-538-al0&t3=2001-01-15&t4=1>)

Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 605.

FEATURES

source
1..606
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="RT0067"

/note="Organ: Kidney tumor; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 32.7%; Score 535; DB 12; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 AAGGAGATCACGTTCTGATGAGACATGATGATCTGAGTACCCAGAGAGAGCTG 567
DB 41 AAGGAGATCACGTTCTGATGAGACATGATGATCTGAGTACCCAGAGAGAGCTG 100

QY 568 GTGTCTGTGCAAGAGATGCTCAACTGCTGGAGGAGCACCAGAAATTCAGAGCAG 627
DB 101 GTGTCTGTGCAAGAGATGCTGAACTGCTGGAGGAGCACCAGAAATTCAGAGCAG 160

QY 628 ATGAAGTCTCTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAGGACCACTGCGCGGT 687
DB 161 ATGAAGTCTCTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAGGACCACTGCGCGGT 220

QY 688 GAGCAGCAGAGCGCTCTGTCGCGCAGCAGCTTGAAGCCATGCTGCGTGAAGTGCAG 747
DB 221 GAGCAGCAGAGCGCTCTGTCGCGCAGCAGCTTGAAGCCATGCTGCGTGAAGTGCAG 280

QY 748 CGGCACACCGCTCCCTCAAGAGAAAGTGTGCAAGCGGCGCCGAGAGAGAGAGAG 807
DB 281 CGGCACACCGCTCCCTCAAGAGAAAGTGTGCAAGCGGCGCCGAGAGAGAGAGAG 340

QY 808 CGCAGAGAGGTGACCTCCGACCTCCAGGTGACACTGAATGACATTCAGCTCAGATGGA 867
DB 341 CGCAGAGAGGTGACCTCCGACCTCCAGGTGACACTGAATGACATTCAGCTCAGATGGA 400

QY 868 CAGCAATGAGCGCAACTCCAAAGTGGCCCAAGAGAAATCGAGCTCGCTGAGAGGCTC 927
DB 401 CAGCAATGAGCGCAACTCCAAAGTGGCCCAAGAGAAATCGAGCTCGCTGAGAGGCTC 460

QY 928 AAGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATATCGCAAAAGTCTTCAACAC 987

DB 461 AAGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATATCGCAAAAGTCTTCAACAC 520
QY 988 AAGGACCTTACAGCAGCAGCTGGTGGATGCGCAAGCTCCAGCAGGCCAGGAGATGC 1042
DB 521 AAGGACCTTACAGCAGCAGCTGGTGGATGCGCAAGCTCCAGCAGGCCAGGAGATGC 575

RESULT 3
BF829485
LOCUS
DEFINITION MR2-HN0034-211200-001-g08 HN0034 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF829485
VERSION BF829485.1
GI:12175105
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
MEDLINE
PUBMED

20202663
10737800

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HN0034-211200-001-g08&t3=2000-12-21&t4=1>)

Seq primer: puc 18 forward
High quality sequence stop: 570.
Location/Qualifiers
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HN0034"

/note="Organ: head normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 29.6%; Score 485; DB 10; Length 570;
Best Local Similarity 99.8%; Pred. No. 8.3e-209;
Matches 535; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 610 CGGAATTCAGAGCAGATGAAGCTCTTACAGAAAAGCAGAGCCAGCTGGTGCAGAG 669
DB 11 CGGAATTCAGAGCAGATGAAGCTCTTACAGAAAAGCAGAGCCAGCTGGTGCAGAG 70

QY 670 AAGGACCACTCGCGGTGAGCAGCAGAGCGCGTCTCTGCGCGCAGCAAGCTTGAAGC 729
DB 71 AAGGACCACTCGCGGTGAGCAGCAGAGCGCGTCTCTGCGCGCAGCAAGCTTGAAGC 130

Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: InCyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1712 row: f column: 21
High quality sequence stop: 766.
Location/Qualifiers
1. .817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:485838"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES
source

Query Match	27.9%	Score 457;	DB:12;	Length 817;
Best Local Similarity	99.4%	Pred. No. 4.4e-196;		
Matches 657;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	1	ATGAAGACCAACAGCAAAAGAACCGGGCTGCCAAATCCAAATCCAAAGAGCCCA 60		
Db	15	ATGAAGACCAACAGCAAAAGAACCGGGCTGCCAAATCCAAATCCAAAGAGCCCA 74		
QY	61	GGACACCGGACAGGACCGGAGGAGCCAGGAGGCGCCAGCCAGCGGCTCTGCA 120		
Db	75	GGACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134		
QY	121	GTAGAAGCAGAGGTCGCGGACAGCAGCGCTCTCGGAAGCGCGGAGGAGTCAAGCC 180		
Db	135	GTAGAAGCAGAGGTCGCGGACAGCAGCGCTCTCGGAAGCGCGGAGGAGTCAAGCC 194		
QY	181	AGAACGGCTCAGTCTGGGGCCCTTGTGTATGTCTGTAGAGGCTGAGCGCCCACTGGAA 240		
Db	195	AGAACGGCTCAGTCTGGGGCCCTTGTGTATGTCTGTAGAGGCTGAGCGCCCACTGGAA 254		
QY	241	GACATCTGAGCACAATCTGTGTGGACATAACACGAGGCGGCGCGGAGGAGTGGGCA 300		
Db	255	GACATCTGAGCACAATCTGTGTGGACATAACACGAGGCGGCGGAGGAGTGGGCA 314		
QY	301	CAGGTCGAGCGGCTGAACCCGAGATGTCAGAGAAGTCCCGGACCTATGTGGCAAGGAT 360		
Db	315	CAGGTCGAGCGGCTGAACCCGAGATGTCAGAGAAGTCCCGGACCTATGTGGCAAGGAT 374		
QY	361	GGGAGGCTGAACCAACTCCAGTAGTCAATGGAGAGGAGGAGGAGGAGGAGGAGTCCA 420		
Db	375	GGGAGGCTGAACCAACTCCAGTAGTCAATGGAGAGGAGGAGGAGGAGGAGGAGTCCA 434		
QY	421	AACACAGAGAGATCCCGGACAGATGACGAGGTCCGAGACCGAGACCATCGAAGCCACAG 480		
Db	435	AACACAGAGAGATCCCGGACAGATGACGAGGTCCGAGACCGAGACCATCGAAGCCACAG 494		
QY	481	GAGAAGAAAAAGCAAGGGTTGGGTAAAGAGATCACTGTCTGTATGTCAGACATTGAAT 540		
Db	495	GAGAAGAAAAAGCAAGGGTTGGGTAAAGAGATCACTGTCTGTATGTCAGACATTGAAT 554		
QY	541	ACTCTGAGTACCCAGAGGAGAGCTGGCTGCTGTGTGCAAGAGATGATGCTGAACCTGCTG 600		
Db	555	ACTCTGAGTACCCAGAGGAGAGCTGGCTGCTGTGTGCAAGAGATGATGCTGAACCTGCTG 614		

ORIGIN

QY 601 GAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCTCTACAGAAAAAGCAGAGCCAGCTG 660
Db 615 GAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCTCTACAGAAAAAGCAGAGCCAGCTG 674
QY 661 G 661
Db 675 G 675
RESULT 6
CD655044 835 bp mRNA linear EST 18-JUN-2003
LOCUS AGENCOURT_14551870 NIA Human HI Embryonic Stem Cell cDNA Library
DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30428424 5', mRNA sequence.
ACCESSION CD655044
VERSION CD655044.1 GI:31894941
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 835)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAMS17 row: h column: 01
High quality sequence stop: 692.
Location/Qualifiers
1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30428424"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human HI Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from Wicell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/Gibco) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pACAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by

FEATURES
source

1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30428424"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human HI Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from Wicell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/Gibco) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pACAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 890)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM13502 row: i column: 18
 High quality sequence stop: 643.
 Location/Qualifiers
 1..890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6157145"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC.72"
 /note="Organ: skin; Vector: pCW-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match	23.1%;	Score 379;	DB 13;	Length 890;
Best Local Similarity	99.8%;	Pred. No. 1.1e-160;		
Matches	429;	Conservative	0;	Mismatches 1;
			Indels	0;
			Gaps	0;

QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGAGCTGAGCGC 231

Db 346 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGAGCTGAGCGC 405

QY 232 CAACCTGGAAGACATACCTAGCAGACATACCTGTGTGGACAATACCAAGGGGGCCCCGGCGAG 291

Db 406 CAACCTGGAAGACATACCTAGCAGACATACCTGTGTGGACAATACCAAGGGGGCCCCGGCGAG 465

QY 292 GATGGGGCAGAGGTGAGCGCGCTGAACCCGAGATGCGAGAGAGTCCCGAGACCTATGTG 351

Db 466 GATGGGGCAGAGGTGAGCGCGCTGAACCCGAGATGCGAGAGAGTCCCGAGACCTATGTG 525

QY 352 GCAAGGAATGGGAGCGCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCTCCCAAG 411

Db 526 GCAAGGAATGGGAGCGCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCTCCCAAG 585

QY 412 GGGGATCCAAACACAGAGAGATCCGGCAGAGTCAAGAGTCCGAGACCGAGACCATCGA 471

Db 586 GGGGATCCAAACACAGAGAGATCCGGCAGAGTCAAGAGTCCGAGACCGAGACCATCGA 645

QY 472 AGGCCACAGGAGAGAAAAGCCAGGGTTGGTAAAGGAGATCAGTGTCTGATCGAG 531

Db 646 AGGCCACAGGAGAGAAAAGCCAGGGTTGGGAAAGGAGATCAGTGTCTGATCGAG 705

QY 532 ACATTGAATCTCTCAGTACCCACAGAGAGAGTGGCTGCTCTCTGCAAGAGATGCT 591

Db 706 ACATTGAATCTCTGAGTACCCACAGAGAGAGTGGCTGCTCTCTGCAAGAGATGCT 765

QY 592 GAACGTGCTGG 601

Db 766 GAACGTGCTGG 775

RESULT 9
 EX370444 LOCUS
 DEFINITION BX370444 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone C50DX06Y023 5'-PRIME, mRNA sequence. linear EST 08-MAY-2003

ACCESSION	EX370444
VERSION	EST.
KEYWORDS	EST.
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 970) Li W.B., Gruber C., Jesses, J. and Polayes, D. Full-length cDNA libraries and normalization: Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 880.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSOBAG046ZA04_CS04345_1custer=880.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOBAG046ZA04_CS04345_1. Location/Qualifiers 1. 970 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODK06YO23" /cell_type="HELA CELLS COT 25-NORMALIZED" /cell_line="HELA" /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES	
source	
ORIGIN	
Query Match	22.9%; Score 375; DB 13; Length 970;
Best Local Similarity	100.0%; Pred. No. 7.4e-159;
Matches	375; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	508 AGGAGATCATCGTTGCTGATCCAGACATTTGAATTACTCTGAGTAGCCCCCAGGAGAACTGG 567
Dd	152 AAGGAGATCATCGTTGCTGATCCAGACATTTGAATTACTCTGAGTAGCCCCCAGGAGAACTGG 211
Qy	568 GCTGTCTCTGTGCAGAAGTATGCTCAACTGCTGGAGGAGCACCGGAATTCACAGAACGAC 627
Dd	212 GCTGTCTCTGTGCAGAAGTATGCTCAACTGCTGGAGGAGCACCGGAATTCACAGAACGAC 271
Qy	628 ATGAGCTCTCTACAGAAAACACAGACGAGCTGGTGACAGAGAAGACACCTGCCCGGT 687
Dd	272 ATGAGCTCTCTACAGAAAACACAGACGAGCTGGTGACAGAGAAGACACCTGCCCGGT 331
Qy	688 GAGCACAGAACGGCCGCTCTCCGCCCGCAGACAAGCTTGAGAGCCCTATGCCCTGAGCTGCAC 747
Dd	332 GAGCACAGAACGGCCGCTCTCCGCCCGCAGACAAGCTTGAGAGCCCTATGCCCTGAGCTGCAC 391
Qy	748 CGGCAACACCGCTCCCTCAAGAGAAAGAGTGTGCAGCGGCCCGGGAGGAGGAGGAAG 807
Dd	392 CGGCAACACCGCTCCCTCAAGAGAAAGAGTGTGCAGCGGCCCGGGAGGAGGAGGAAG 451
Qy	808 CGCAAGAGGTGACCTCCGCACTTCAGGTGCACACTGAATGACATTCAGTGCAGATGGAA 867
Dd	452 CGCAGGAGGTGACCTCCGCACTTCAGGTGCACACTGAATGACATTCAGTGCAGATGGAA 511
Qy	868 CAGCACAAATGAGCGC 882
Dd	512 CAGCACAAATGAGCGC 526

RESULT 10
BM800456

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Bd	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
LOCUS	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
DEFINITION	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
ACCESSION	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
VERSION	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
KEYWORDS	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
SOURCE	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
ORGANISM	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
REFERENCE	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
AUTHORS	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
TITLE	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
JOURNAL	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
COMMENT	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
FEATURES	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
ORIGIN	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Query Match	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Best Local Similarity	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Matches	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Conservative	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Mismatches	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Indels	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
Gaps	BX362957	1195 bp	mRNA	linear	EST 05-MAY-2003			
1	ATGAAGAACCAAGACCAAAAGAACCGGGCTGCAAAACAAATCCAAATCCAAAGACGAGCCCA	60						
120	ATGAAGAACCAAGACCAAAAGAACCGGGCTGCAAAACAAATCCAAATCCAAAGACGAGCCCA	179						
61	GGACAAACCGGAAGCAGGACCCGAGGAGCCGACGAGGCGCCGACGAGCGGCTCTCTGCA	120						
180	GGACAAACCGGAAGCAGGACCCGAGGAGCCGACGAGGCGCCGACGAGCGGCTCTCTGCA	239						
121	GTAGAGCAGAGAGGTCCGCGAGCAGCAGCGTCTCTCGAGAGCCGAGGAGGTCTCAAGCC	180						
240	GTAGAGCAGAGAGGTCCGCGAGCAGCAGCGTCTCTCGAGAGCCGAGGAGGTCTCAAGCC	299						
181	AGAACGGCTCAGTCTGGGSCCTTCTCGTATGTCTCTGAGGAGTCTGAGCCGCACTGGAA	240						
300	AGAACGGCTCAGTCTGGGSCCTTCTCGTATGTCTCTGAGGAGTCTGAGCCGCACTGGAA	359						
241	GACATCTGACACATCTGTGTGACAAATACCGAGGAGGAGGAGGAGGAGGAGGAGGAGG	300						
360	GACATCTGACACATCTGTGTGACAAATACCGAGGAGGAGGAGGAGGAGGAGGAGGAGG	419						
301	CAGGGGTGAGCCGGCTGAACCCGAGAGTGCAGAGAGTCCCGGAGCTATGTGG	352						

```

317 AACCCGAAGATGCAGAGAATCCCGGACCTATTGTGGCAAGGAATGGGAGCCTGAACCAA 376
333 AACCCGAAGATGCAGAGAATCCCGGACCTATTGTGGCAAGGAATGGGAGCCTGAACCAA 392
377 CTCAGTAGTCAATGGAGAGAGAAACCTCCCAAGGGGGATCCAAACACAGAGAGATCC 436
393 CTCAGTAGTCAATGGAGAGAGAGAAACCTCCCAAGGGGGATCCAAACACAGAGAGATCC 452
437 GGCAGAGTGCAGAGGTGG 455
453 GGCAGAGTGCAGAGGTGG 471

RESULT 14
BF796069 965 bp mRNA linear EST 12-JAN-2001
LOCUS 602258906F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342392 5',
DEFINITION mRNA sequence.
ACCESSION BF796069
VERSION BF796069.1 GI:12101123
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9957 row: k column: 01
High quality sequence stop: 557.
FEATURES
source
1..965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4342392"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.7%; Score 241; DB 10; Length 965;
Best Local Similarity 100.0%; Pred. No. 4.9e-98;
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 838 ACATGATGATTCAGCTGCAGATGGAACAGCAGCAATGAGGCGCACTCCAGCTGCGC 897
DB 16 ACATGATGATTCAGCTGCAGATGGAACAGCAGCAATGAGGCGCACTCCAGCTGCGC 75
QY CAAGAGAACATGGAGCTGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGC 957
DB 76 CAAGAGAACATGGAGCTGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGC 135
QY GAGGAGCATATCGCAAAAGTCTTCAAAACAGAGGACCTACAAACAGCAGCTGGTGGATGCC 1017
DB 136 GAGGAGCATATCGCAAAAGTCTTCAAAACAGAGGACCTACAAACAGCAGCTGGTGGATGCC 195

```

```

QY 1018 AAGCTCCAGCAGCGCCAGGAGATGCTAAAGAGGAGGAGGAGCGGACCGGAGGAG 1077
DB 196 AAGCTCCAGCAGCGCCAGGAGATGCTAAAGAGGAGGAGGAGCGGACCGGAGGAG 255
QY 1078 G 1078
DB 256 G 256

RESULT 15
BI055054/c
LOCUS CM3-GN0092-050201-707-f12 GN0092 Homo sapiens cDNA, mRNA sequence.
DEFINITION CM3-GN0092-050201-707-f12 GN0092 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI055054
VERSION BI055054.1 GI:14462584
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 281)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-GN0092-
050201-707-f12&t3=2001-02-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 281.
Location/Qualifiers
1..281
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0092"
/notes="Organ: placenta_normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN
Query Match 14.1%; Score 231; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.5e-93;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1000 CAGCAGCTGGTGGATGCCAAGCTCCAGAGCCCGAGGAGATGCTAAAGGAGGAGAG 1059
DB 281 CAGCAGCTGGTGGATGCCAAGCTCCAGAGCCCGAGGAGATGCTAAAGGAGGAGAG 222
QY 1060 CGGCACACGCGGAGAGGAGGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1119

```

```

Db      221  CGGCACCGCGGAGAGGATTTCTCTCTGAAAGAGGACGTAGAGTCCACAGGAGTGTGT 162
QY      1120  GAGCTGATCAAGCAGCAGACAGACCCACCTGAGCAGACAGCTTGCCCTATACACAGAGAG 1179
Db      161  GAGCTGATGAAGCAGCAGACAGACCCACCTGAGCAGACAGCTTGCCCTATACACAGAGAG 102
QY      1180  TTGTGAGAGTTCAGAACACACATTTTCCAAAAGCAGGAGGTATTCCACCACA 1230
Db      101  TTGTGAGAGTTCAGAACACACATTTTCCAAAAGCAGGAGGTATTCCACCACA 51

RESULT 16
BF915348      338 bp      mRNA      linear      EST 18-JAN-2001
LOCUS      MR3-UT0106-041200-005-a12 UT0106 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BF915348
ACCESSION      BF915348
VERSION      BF915348.1 GI:12306806
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 338)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J., and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR3&t2=MR3-UT0106-
041200-005-a12&t3=2000-12-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 338.

FEATURES      source
1..338
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UT0106"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 136,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      12.6%; Score 206; DB 10; Length 338;
Best Local Similarity 99.6%; Pred. No. 3.4e-82;
Matches 256; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      94  GAGCGCCCGCAGCGCGGCTCTGTCAGTAGAGCAGAGTCCCGCAGCAGCCAGGCT 153
Db      82  GAGCGCCCGCAGCGCGGCTCTGTCAGTAGAGCAGAGTCCCGCAGCAGCCAGGCT 141

```

```

QY      154  CCTCGAAGCGCGGAGGCTGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTC 213
Db      142  CCTCGAAGCGCGGAGGCTGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTC 201
QY      214  TCTGAGAGCTGAGCCGCCCACTGGAAGACATCTGAGCACAATCTGTGTGGACAATAAC 273
Db      202  TCTGAGAGCTGAGCCGCCCACTGGAAGACATCTGAGCACAATCTGTGTGGACAATAAC 261
QY      274  CAGGGGGCGCCCGCGGAGGATGGGGCAGCGGCTGAGCCGGCTGAACCCGAAGATGCGAG 333
Db      262  CAGGGGGCGCCCGCGGAGGATGGGGCAGCGGCTGAGCCGGCTGAACCCGAAGATGCGAG 321
QY      334  AAGTCCCGGACCTATGT 350
Db      322  AAGTCCCGGACCTATGT 338

RESULT 17
BF850082      316 bp      mRNA      linear      EST 16-JAN-2001
LOCUS      RCO-EN0080-161100-021-a03 EN0080 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BF850082
ACCESSION      BF850082
VERSION      BF850082.1 GI:12237232
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 316)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J., and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RCO&t2=RCO-EN0080-
161100-021-a03&t3=2000-11-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 316.

FEATURES      source
1..316
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0080"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      10.9%; Score 178; DB 10; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;

```

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 511 GAGATACGTTGCTGATGACACATTGAAATCTCTAGTACCCAGAGGAGAGCTGGCT 570
 Db 123 GAGATACGTTGCTGATGACACATTGAAATCTCTAGTACCCAGAGGAGAGCTGGCT 182
 QY 571 GCTCTGTGCAAGAAAGTATGCTGAACCTCTCGAGAGACCCGGAATTACAGAGACAGATG 630
 Db 183 GCTCTGTGCAAGAAAGTATGCTGAACCTCTCGAGAGACCCGGAATTACAGAGACAGATG 242
 QY 631 AAGCTCTACAGAAAGACGAGCCAGCTGCTGCAAGAGAGACCACTGCGCGGTG 688
 Db 243 AAGCTCTACAGAAAGACGAGCCAGCTGCTGCAAGAGAGAGACCACTGCGCGGTG 300

RESULT 18
 AA423995/c
 LOCUS AA423995 525 bp mRNA linear EST 16-OCT-1997
 DEFINITION zv79h08.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
 IMAGE:759903 3', mRNA sequence.
 AA423995
 ACCESSION AA423995
 VERSION AA423995.1 GI:2102974
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 525)
 AUTHORS Hallier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 TITLE Washu-Merck EST Project 1997
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41m13 fwd. RT from Amersham.
 High quality sequence stop: 400.

FEATURES
 Location/Qualifiers
 1..525
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IWAG:759903"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares total fetus Nb2HF8_9w"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTAGTGGAGCGGCTTAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 8.8%; Score 145; DB 9; Length 525;
 Best Local Similarity 100.0%; Pred. No. 1.7e-54;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GGGCTGCAACCAATCCAAATCCAAAAGCAGCCAGCAACCGGAAGCAGGACCCGAGG 85
 Db 506 GGGCTGCAACCAATCCAAATCCAAAAGCAGCCAGCAACCGGAAGCAGGACCCGAGG 447

QY 86 GAGCCAGAGAGCGGCCAGCCAGCGGCTCTGCAGTAGAAGCAGAGGTCCCGGACGA 145
 Db 446 GAGCCAGAGAGCGGCCAGCCAGCGGCTCTGCAGTAGAAGCAGAGGTCCCGGACGA 387
 QY 146 GCCAGGCTCTCGGAGCGGAGGG 170
 Db 386 GCCAGGCTCTCGGAGCGGAGGG 362

RESULT 19
 BQ310163
 LOCUS BQ310163 297 bp mRNA linear EST 16-MAY-2002
 DEFINITION MR0-BT3503-140501-001-c02 BT3503 Homo sapiens cDNA, mRNA sequence.
 BQ310163
 ACCESSION BQ310163
 VERSION BQ310163.1 GI:20852706
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 297)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20302663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR0&t2=MR0-BT3503-
 140501-001-c02&t3=2001-05-14&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 17
 High quality sequence stop: 198.

FEATURES
 Location/Qualifiers
 1..297
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BT3503"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

ORIGIN
 Query Match 8.4%; Score 138; DB 13; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.5e-51;
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CAACAATCCAAATCCAAAAGCAGCCAGCAACCGGAAGCAGGACCCGAGGAGCCCA 92
 Db 71 CAACAATCCAAATCCAAAAGCAGCCAGCAACCGGAAGCAGGACCCGAGGAGCCCA 130
 QY 93 GGAGCGGCCAGCGCGGCTCTGCAGTAGAAGCAGAGGTCCCGGACGAGCCAGGC 152


```

Db      131 GGAGCGCGCCAGCGCGCTCTCTCAGTAGAAGCAGAAGTCCCGGAGCAGCGCAGG 190
QY      153 TCCTCGAGAGCCGAGGG 170
Db      191 TCCTCGAAGCGGAGGG 208

RESULT 20
BQ310185
LOCUS      268 bp mRNA linear EST 16-MAY-2002
DEFINITION MR0-BT3503-180501-002-a08 BT3503 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BQ310185
VERSION     BQ310185.1 GI:20852728
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 268)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-270701
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR0&t2=MR0-BT3503-
180501-002-a08&t3=2001-05-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 268.
FEATURES
source
Location/Qualifiers
1..268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT3503"
/notes="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 8.1%; Score 133; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 4.6e-49;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      172 GCTCAAGCCAGACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTCAGAGAGCTGACCGC 231
Db      136 GCTCAAGCCAGACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTCAGAGAGCTGACCGC 195
QY      232 CAACTGGAACATCTAGACATATCTGTGTGACAAATAACAGGGGGCCCGCCGCGAG 291
Db      196 CAACTGGAACATCTAGACATATCTGTGTGACAAATAACAGGGGGCCCGCCGCGAG 255
QY      292 GATGGGGCACAGG 304

```

```

Db      256 GATGGGGCACAGG 268

RESULT 21
W07246
LOCUS      605 bp mRNA linear EST 25-APR-1996
DEFINITION za94g06.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone
IMAGE:300250 5', mRNA sequence.
ACCESSION  W07246
VERSION     W07246.1 GI:1281248
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 605)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 344.
FEATURES
source
Location/Qualifiers
1..605
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1245174"
/db_xref="taxon:9606"
/clone="IMAGE:300250"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal lung NbHL19W"
/notes="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Ronaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHL19W."
ORIGIN
Query Match 7.9%; Score 130; DB 14; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-47;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAAGAACCAAGACAAAGAACCGGGCTGCCAACAATCCAAATCCAAAAGCAGCCCA 60
Db      54 ATGAAGAACCAAGACAAAGAACCGGGCTGCCAACAATCCAAATCCAAAAGCAGCCCA 113
QY      61 GGACAAACCGAAGCAGGACCCGAGGAGCCCGAGGCGGCCAGCGCGGCTCTCTGCA 120
Db      114 GGACAAACCGAAGCAGGACCCGAGGAGCCCGAGGCGGCCAGCGCGGCTCTCTGCA 173
QY      121 GTAGAGCAG 130
Db      174 GTAGAGCAG 183

```

RESULT 22
AA423994
LOCUS
DEFINITION
zv79h08.r1 Soares_total_fetus Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:759903 5', mRNA sequence.

ACCESSION
AA423994
VERSION
AA423994.1 GI:2102973
KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 443)
Hallier L., Allen M., Bowles L., Dubuque T., Geisel G., Jost S.,
Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin J.,
Moore B., Schellenberg K., Scheppoe M., Tan F., Theising B.,
White Y., Wylie T., Waterston R. and Wilson R.
WashU-Merck EST Project 1997

TITLE
Unpublished (1997)

JOURNAL
Contact: Wilson RK

COMMENT
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28n13 rev2 ET from Amersham
High quality sequence stop: 432.

FEATURES
Location/Qualifiers
1..443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:759903"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares_total_fetus Nb2HF8_9w"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 7.1%; Score 116; DB 9; Length 443;
Best Local Similarity 100.0%; Pred. No. 2.5e-41;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGAACCAAGACAAAGAGCGGGCTGCCAAATCCCAATCCAAAGACGACCCA 60
DB 59 ATGAGAACCAAGACAAAGAGCGGGCTGCCAAATCCCAATCCAAAGACGACCCA 118
QY 61 GGACAAACCGAAGCAGACCCGAGGAGCCAGAGCGGCCCGCCAGCGGGCTCC 116
DB 119 GGACAAACCGAAGCAGACCCGAGGAGCCAGAGCGGCCCGCCAGCGGGCTCC 174

RESULT 23
AI368649/c
LOCUS
DEFINITION
qv60g11.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2016452 3',
mRNA sequence.

ACCESSION
AI368649
VERSION
AI368649.1 GI:4147402
KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 388)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGA), Tumor Gene Index
Unpublished (1998)

JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.

COMMENT
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 718 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 317.

FEATURES
Location/Qualifiers
1..388
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2016452"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/notes="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 6.8%; Score 112; DB 9; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.6e-39;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 569 CTGCTCTCTGCAAGAAGTATCTGAATCTGCTGGAGGAGCACCAGGATTCACAGAGCAGA 628
DB 357 CTGCTCTCTGCAAGAAGTATCTGAATCTGCTGGAGGAGCACCAGGATTCACAGAGCAGA 298

QY 629 TCAAGCTCTACAGAAAAGCAGAGCCAGCTGCTGCAGAGAGGACCACT 680
DB 297 TCAAGCTCTACAGAAAAGCAGAGCCAGCTGCTGCAGAGAGGACCACT 246

RESULT 24
BG015280
LOCUS
DEFINITION
RC2-GN0359-251200-013-f05 GN0359 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG015280
VERSION
BG015280.1 GI:12467327
KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 280)
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Costa F.F.,
Goldman G.H., Carvalho A.F., Matekuma A., Baia G.S., Simpson D.H.,
Brunstein A., deOliveira P.S., Bucher P., Jongeneel C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL
 MEDLINE
 20202663
 PUBMED
 10737800
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC2&t2=RC2-GN0359-251200-013-f05&t3=2000-12-25&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 6
 High quality sequence stop: 280.
 Location/Qualifiers
 1. .280
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="GN0359"
 /note="Organ: placenta normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 ORIGIN
 Query Match 6.7%; Score 109; DB 10; Length 280;
 Best Local Similarity 100.0%; Pred. No. 3.6e-38;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1267 AGAAGCTGGAGAAACCAACCATGTCACCGTCCCGTGGAGAGCAGCAGGCGC 1326
 Db 133 AGAAGCTGGAGAAACCAACCATGTCACCGTCCCGTGGAGAGCAGCAGGCGC 192
 QY 1327 CTGCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGG 1375
 Db 193 CTGCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGG 241
 RESULT 25
 BF081747
 LOCUS
 DEFINITION
 PM2-AN0089-050900-004-a09 AN0089 Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 BF081747
 VERSION
 BF081747.1 GI:10875577
 KEYWORDS
 EST.
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 124)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL
 MEDLINE
 20202663
 TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL
 MEDLINE
 20202663

10737800
 PUBMED
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC2&t2=RC2-AN0089-050900-004-a09&t3=2000-09-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 123.
 Location/Qualifiers
 1. .124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="AN0089"
 /note="Organ: amnion normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 ORIGIN
 Query Match 6.1%; Score 100; DB 10; Length 124;
 Best Local Similarity 100.0%; Pred. No. 4e-34;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 869 AGCACATGAGCGCAACTCCCAAGCTGCGCCAGAGACATGAGCTGCTGAGAGGCTCA 928
 Db 11 AGCACATGAGCGCAACTCCCAAGCTGCGCCAGAGACATGAGCTGCTGAGAGGCTCA 70
 QY 929 AGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATAT 968
 Db 71 AGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATAT 110
 RESULT 26
 A1468002/c
 LOCUS
 DEFINITION
 tJ84f05.x1 Soares NSF P8 9W OT PA P S1 Homo sapiens cDNA clone IMAGE:2148225 3', similar to SW:1114 HUMAN P40222 INTERLEUKIN-14 PRECURSOR ; contains element MER32 PTF5 repetitive element ;, mRNA sequence.
 ACCESSION
 A1468002
 VERSION
 A1468002.1 GI:4330092
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 414)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgsabbs-remail.nih.gov
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 632 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 112.
 Location/Qualifiers
 1. .414
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2148225"
 /lab_host="DH10B"
 /clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NBHP pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HF8-9W pool 1: 758280-760589, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 5.9%; Score 97; DB 9; Length 414;
 Best Local Similarity 100.0%; Pred. No. 16-32; Indels 0; Gaps 0;
 Matches 97; Conservative 0; Mismatches 0;

QY 1155 ACAGCTTGCCTATACACAGAGAGTTGAGGAGTTCACAGACACATTTCCAAAAGCAG 1214
 Db 414 ACAGCTTGCCTATACACAGAGAGTTGAGGAGTTCACAGACACATTTCCAAAAGCAG 355

QY 1215 CGAGGTATTCACACATTCAGCAGGAGATGGAAG 1251
 Db 354 CGAGGTATTCACACATTCAGCAGGAGATGGAAG 318

RESULT 27
 AQ892680 850 bp DNA linear GSS 10-NOV-1999
 LOCUS HS 3143_A1_Q05_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3143 Col=9 Row=M, genomic survey sequence.
 ACCESSION AQ892680
 VERSION AQ892680.1 GI:6349870
 KEYWORDS GSS:
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 850)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 9380589
 PUBMED 1049764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 3143 row: M column: 9
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 850.
 Location/Qualifiers

1. 850
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=3143 Col=9 Row=M"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

ORIGIN
 Query Match 5.5%; Score 90; DB 28; Length 850;
 Best Local Similarity 100.0%; Pred. No. 1.7e-29; Indels 0; Gaps 0;
 Matches 90; Conservative 0; Mismatches 0;

QY 508 AAGGAGATCATGTTGCTGATGACAGACATTGAATACTCTGAGTACCCAGAGAGAGCTG 567
 Db 337 AAGGAGATCATGTTGCTGATGACAGACATTGAATACTCTGAGTACCCAGAGAGAGCTG 396

QY 568 GCTGCTCTGTGCAAGAAGTATGCTGAACCTG 597
 Db 397 GCTGCTCTGTGCAAGAAGTATGCTGAACCTG 426

RESULT 28
 AA761241 254 bp mRNA linear EST 07-FEB-1998
 LOCUS AA761241/c
 DEFINITION nr220f05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1288353 3', mRNA sequence.
 ACCESSION AA761241
 VERSION AA761241.1 GI:2810171
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 254)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1280 Std Error: 0.00
 Seq primer: -40ml3 fwd. EF from Amersham
 High quality sequence stop: 247.
 Location/Qualifiers

FEATURES

1. 254
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1288353"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP GCB1"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAGTGGAGCGGCTCATTTTTTTTTTTT-3',]. Double-stranded cDNA was ligated to Eco RI adaptors

FEATURES

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pMT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 4.9%; Score 81; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.8e-25;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1012 GATGCCAAGTCCAGCAGGCCAGGATGCTTAAGAGGCGAGAGCGGACACAGCGG 1071
Db 254 GATGCCAAGTCCAGCAGGCCAGGATGCTTAAGAGGCGAGAGCGGACACAGCGG 195
QY 1072 GGAAGGATTTCTCTGAAA 1092
Db 194 GAGAGGATTTCTCTGAAA 174

RESULT 29
BF765947/c
LOCUS BF765947 223 bp mRNA linear EST 12-JAN-2001
DEFINITION CM3-CS0044-111100-482-h12 CS0044 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF765947
VERSION BF765947.1 GI:112113847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 223)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM3&t2=CM3-CS0044-111100-482-h12&t3=2000-11-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 223.
Location/Qualifiers
1..223
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CS0044"
/note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source
1..223
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CS0044"
/note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 4.8%; Score 79; DB 10; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-24;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1059 CGGCACACGCGGAGAGGATTTCTCTGAAAAGAGCAGTAGATCCACAGAGTGTG 1118
Db 135 CGGCACACGCGGAGAGGATTTCTCTGAAAAGAGCAGTAGATCCACAGAGTGTG 76
QY 1119 TGAGCTGATGAAGCAGCAA 1137
Db 75 TGAGCTGATGAAGCAGCAA 57

RESULT 30
BF753862
LOCUS BF753862 273 bp mRNA linear EST 10-JAN-2001
DEFINITION RC1-CT0045-041000-011-h07 CT0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF753862
VERSION BF753862.1 GI:12080538
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 273)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC1&t2=RC1-CT0045-041000-011-h07&t3=2000-10-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 273.
Location/Qualifiers
1..273
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0045"
/note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC1&t2=RC1-CT0045-041000-011-h07&t3=2000-10-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 273.
Location/Qualifiers
1..273
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0045"
/note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source
1..273
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0045"
/note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 4.0%; Score 65; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1571 ACCGAGGACCGGAGCAGAGCAGATCAGGCCACTGGCCCTCAAGAGCCACTCGG 1630

```

Db      43  ACCGAGGACCGGACGACGAGCATCAGGCCAGACTGGCGCTCAAGAGCCCACTCCG 102
QY      1631  CCAGG 1635
Db      103  CCAGG 107

RESULT 31
LOCUS   BE932678/c
DEFINITION RC3-H70860-170800-011-a08 HT0860 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE932678
VERSION   BE932678.1 GI:10458754
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 429)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3-HT0860-170
          800-011-a08&t3=2000-08-17&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 429.
          Location/Qualifiers
          1..429
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="HT0860"
             /note="Organ: head neck; Vector: puc18; Site_1: SmaI;
             Site_2: SmaI; A mini-library was made by cloning products
             derived from ORESTES PCR (U.S. Letters Patent application
             No. 196,716 - Ludwig Institute for Cancer Research)
             profiles into the pUC 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."

ORIGIN
Query Match 3.8%; Score 63; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.8e-17;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286  GCGGAGATGGGCACAGGTGAGCGGTGAACCGAGATCGAGAGTCCCGGACC 345
Db      132  GCGGAGATGGGCACAGGTGAGCGGTGAACCGAGATCGAGAGTCCCGGACC 73

QY      346  TAT 348
Db      72  TAT 70

RESULT 32
LOCUS   BQ310186
DEFINITION MRO-BT3503-180501-002-h12 BT3503 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ310186
VERSION   BQ310186.1 GI:20852729
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 152)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MRO-BT3503-
          180501-002-h12&t3=2001-05-18&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 152.
          Location/Qualifiers
          1..152
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="BT3503"
             /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
             SmaI; A mini-library was made by cloning products derived
             from ORESTES PCR (U.S. Letters Patent application No.
             196,716 - Ludwig Institute for Cancer Research) profiles
             into the pUC 18 vector. Reverse transcription of tissue
             mRNA and cDNA amplification were performed under low
             stringency conditions."

ORIGIN
Query Match 3.7%; Score 61; DB 13; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      220  GAGCTGAGCCGCCACTGGAGACATCTGAGCACATCTGTGGACATAACACGGGG 279
Db      51  GAGCTGAGCCGCCACTGGAGACATCTGAGCACATCTGTGTGGACATACACGGGG 110

QY      280  G 280
Db      111  G 111

RESULT 33
LOCUS   CB427292
DEFINITION 602860 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION CB427292
VERSION   CB427292.1 GI:29200195
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 152)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MRO-BT3503-
          180501-002-h12&t3=2001-05-18&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 152.
          Location/Qualifiers
          1..152
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="BT3503"
             /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
             SmaI; A mini-library was made by cloning products derived
             from ORESTES PCR (U.S. Letters Patent application No.
             196,716 - Ludwig Institute for Cancer Research) profiles
             into the pUC 18 vector. Reverse transcription of tissue
             mRNA and cDNA amplification were performed under low
             stringency conditions."

ORIGIN
Query Match 3.8%; Score 63; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.8e-17;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286  GCGGAGATGGGCACAGGTGAGCGGTGAACCGAGATCGAGAGTCCCGGACC 345
Db      132  GCGGAGATGGGCACAGGTGAGCGGTGAACCGAGATCGAGAGTCCCGGACC 73

QY      346  TAT 348
Db      72  TAT 70

```

KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovidae; Bos.
1 (bases 1 to 628)
Smith, T.P.L., Roberts, A.J., Echterkamp, S.E., Chitko-McKown, C.G., Wray, J.E. and Keele, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries

TITLE Unpublished (2003)
JOURNAL Contact: Smith TPL
COMMENT USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross match v0.990329.
Plate: FQ8019 row: G column: 7
Seq primer: GTAATACGACTCATATAGGG.
Location/Qualifiers
1..628
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."

FEATURES source
1..628
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/dev_stage="Adult"
/clone_lib="MARC 6BOV"
/note="Organ: lung tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 3.2%; Score 53; DB 14; Length 628;
Best Local Similarity 100.0%; Pred. No. 1e-12;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 754 AACGGTCTCCCTCAAGGAAGGTGTGCAGCGCGCGGAGGAGGAGAA 806
|||||
441 AACGGTCTCCCTCAAGGAAGGTGTGCAGCGCGCGGAGGAGGAGAA 493
|||||

Db

RESULT 34
BI013335/c
LOCUS
DEFINITION PM1-ET0205-220101-004-g02 ET0205 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI013335
VERSION BI013335.1 GI:14417406
KEYWORDS EST:
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 413)

REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM1et2-PM1-ET0205-220101-004-g02&3=2001-01-22&4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 292.
Location/Qualifiers
1..413
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ET0205"
/note="Organ: lung tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES source
1..413
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ET0205"
/note="Organ: lung tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 3.1%; Score 51; DB 12; Length 413;
Best Local Similarity 100.0%; Pred. No. 7.9e-12;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 659 TGGTCAAGAGAAGGACCACTCGCGGTGAGCAGCAGCAAGCGCTCTCTGG 709
|||||
352 TGGTCAAGAGAAGGACCACTCGCGGTGAGCAGCAGCAAGCGCTCTCTGG 302
|||||

Db

RESULT 35
N78816/c
LOCUS
DEFINITION z94g06.sl Soares fetal lung NbHL19W Homo sapiens cDNA clone
IMAGE:300250 3', mRNA sequence.
ACCESSION N78816
VERSION N78816.1 GI:1241517
KEYWORDS EST:
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 414)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevas, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 342.
Location/Qualifiers
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="GDB:1245174"
/db_xref="taxon:9606"
/clone="IMAGE:300250"
/dev_stage="19 weeks"

FEATURES source
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="GDB:1245174"
/db_xref="taxon:9606"
/clone="IMAGE:300250"
/dev_stage="19 weeks"

/lab host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_fetal_lung_NbHL19W"
 /notes="Organ: lung; Vector: p773D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773D vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NbHL19W."

ORIGIN

Query Match 3.1%; Score 50; DB 14; Length 414;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 CAGGAGGTGACCTCGCACTCCAGGTGACATGACATTCAGCTGC 859

Db 309 CAGGAGGTGACCTCGCACTCCAGGTGACATGACATTCAGCTGC 260

RESULT 36
 AI655176/c 427 bp mRNA linear EST 17-DEC-1999
 LOCUS W667f01.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2310745 3',
 DEFINITION mRNA sequence.

ACCESSION AI655176

VERSION AI655176.1 GI:4739155

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 427)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapsb-remail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 529 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..427

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2310745"

/tissue_type="pooled germ cell tumors"

/lab host="DH10B"

/clone_lib="NCI_CGAP GC6"

/notes="Vector: p773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI_CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

"

ORIGIN

Query Match 3.1%; Score 50; DB 9; Length 427;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 GACCGAGACCATCGAAGCCACAGAGAGAAAAAAGCAAGGTTTGGG 506

Db 188 GACCGAGACCATCGAAGCCACAGAGAGAAAAAAGCAAGGTTTGGG 139

RESULT 37

BGI50260/c

LOCUS BGI50260.1

DEFINITION 427 bp mRNA linear EST 05-FEB-2001

mRNA sequence.

ACCESSION BGI50260

VERSION BGI50260.1 GI:12662290

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 427)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapsb-remail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..427

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3442872"

/tissue_type="pooled germ cell tumors"

/lab host="DH10B"

/clone_lib="NCI_CGAP GC6"

/notes="Vector: p773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI_CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

"

ORIGIN

Query Match 3.1%; Score 50; DB 12; Length 427;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 GACCGAGACCATCGAAGCCACAGAGAGAAAAAAGCAAGGTTTGGG 506

Db 188 GACCGAGACCATCGAAGCCACAGAGAGAAAAAAGCAAGGTTTGGG 139

RESULT 38

AI970193/c

LOCUS AI970193

DEFINITION 428 bp mRNA linear EST 08-MAR-2000

Subtraction by Bento Soares and M. Fatima Bonaldo.

"

VERSION
KEYWORDS
SOURCE
ORGANISM

CP742562.1 GI:37638901

EST.
Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 527)

REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 527

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/tissue_type="whole eye"

/dev_stage="embryo 12.5, 13.5, 14.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP HB0"

/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is TTATGTGAAGT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH)."

ORIGIN

Query Match 2.9%; Score 47; DB 14; Length 527;

Best Local Similarity 100.0%; Pred. No. 5.3e-10;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

826 CACTTCCAGGTGACACTGATGACATTCAGCTGCAGATGGAACAGCA 872

Db

426 CACTTCCAGGTGACACTGATGACATTCAGCTGCAGATGGAACAGCA 472

RESULT 41

BU613887

LOCUS

BU613887 539 bp mRNA linear EST 20-FEB-2003

UI-M-EWO-cba-p-14-0-UI.r1 NIH BMAP_EWO Mus musculus cDNA clone

UI-M-EWO-cba-p-14-0-UI 5', mRNA sequence.

BU613887.1 GI:23280102

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 539)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 539

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="UI-M-EWO-cba-p-14-0-UI"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_EWO"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction,

ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA

tail is GTGCTGGGAA. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project

(BMAP): 'Gene Discovery in the Developing Mouse Nervous

System', supported by National Institutes of Mental Health

(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 539;

Best Local Similarity 100.0%; Pred. No. 5.3e-10;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

826 CACTTCCAGGTGACACTGATGACATTCAGCTGCAGATGGAACAGCA 872

Db

386 CACTTCCAGGTGACACTGATGACATTCAGCTGCAGATGGAACAGCA 432

RESULT 42

CD807318

LOCUS

CD807318 551 bp mRNA linear EST 15-JUL-2003

DEFINITION

IMAGE:30542366 5', mRNA sequence.

ACCESSION

CD807318

VERSION

CD807318.1 GI:32466144

KEYWORDS

EST.

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 551)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project

```

(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1. 551
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/cloned="IMAGE:30542366"
/tissue_type="whole eye"
/dev_stage="embryo 15.5,16.5,17.5,18.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP GW0"
/notes="Organ: Eye; Vector: pyx-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaudo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dr primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CTGGGTCCTC. This library was created for the University of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Query Match 2.9%; Score 47; DB 14; Length 551;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGATGACATTCAGTCGAGTGAACAGCA 872
Db 287 CACTTCAGGTGACACTGATGACATTCAGTCGAGTGAACAGCA 333

RESULT 43
AW320794 580 bp mRNA linear EST 25-JAN-2000
LOCUS uc23el0.v1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:2615754 5'
DEFINITION similar to SW:IL14_HUMAN P40222 INTERLEUKIN-14 PRECURSOR ; mRNA
sequence.
ACCESSION AW320794.1 GI:6750338
VERSION AW320794.1
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 580)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

MGI:1023854
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 374.
Location/Qualifiers
1. 580
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/cloned="IMAGE:2615754"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match 2.9%; Score 47; DB 10; Length 580;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGATGACATTCAGTCGAGTGAACAGCA 872
Db 324 CACTTCAGGTGACACTGATGACATTCAGTCGAGTGAACAGCA 370

RESULT 44
AW412004 620 bp mRNA linear EST 08-FEB-2000
LOCUS uc055f09.v1 NCI CGAP Lu29 Mus musculus cDNA clone IMAGE:2646473 5'
DEFINITION similar to SW:IL14_HUMAN P40222 INTERLEUKIN-14 PRECURSOR ; mRNA
sequence.
ACCESSION AW412004
VERSION AW412004.1 GI:6937859
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 620)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

MGI:1026925
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 431.
Location/Qualifiers
1. 620
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/cloned="IMAGE:2646473"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

```

ORIGIN

Query Match 2.9%; Score 47; DB 10; Length 620;
 Best Local Similarity 100.0%; Pred. No. 5.3e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 872
 |||||
 Db 322 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 368

RESULT 45

CD578468
 LOCUS
 DEFINITION UI-M-FYO-cfs-n-14-0-UI.r1 NIH_BMAP_FYO Mus musculus cDNA clone EST 09-JUL-2003
 IMAGE:6855471 5', mRNA sequence.
 CD578468
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 647)
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousef1.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Location/Qualifiers
 1. 647
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6855471"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FYO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGACG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 14; Length 647;
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 872
 |||||
 Db 384 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 430

RESULT 46

BU613886
 LOCUS
 DEFINITION UI-M-EWO-cba-p-12-0-UI.r1 NIH_BMAP_EWO Mus musculus cDNA clone EST 20-FEB-2003
 UI-M-EWO-cba-p-12-0-UI 5', mRNA sequence.
 BU613886
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 653)
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, Bento-Soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Location/Qualifiers
 1. 653
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UI-M-EWO-cba-p-12-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EWO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail is GTGCGTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 653;
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 872
 |||||
 Db 388 CACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACAGCA 434

RESULT 47

BU701604
 LOCUS
 DEFINITION UI-M-FIO-byp-e-23-0-UI.r1 NIH_BMAP_FIO Mus musculus cDNA clone EST 15-JUL-2003
 IMAGE:5721214 5', mRNA sequence.
 BU701604
 ACCESSION
 VERSION

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 699)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
 source
 1..699
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /clone="IMAGE:5721214"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH_BMAP_F10"
 /notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCAGCAGC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 2.9%; Score 47; DB 13; Length 699;
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 826 CACTTCGAGTGACACTGATGACATTCAGCTCGAGATGGAACAGCA 872
 Db 435 CACTTCGAGTGACACTGATGACATTCAGCTCGAGATGGAACAGCA 481

RESULT 48
 CB525790
 LOCUS 710 bp mRNA linear EST 09-JUL-2003
 DEFINITION UI-M-PYO-cfe-f-03-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
 IMAGE:6848356 5', mRNA sequence.
 CB525790
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 710)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)

REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found at <http://genome.uiowa.edu/distribution/mousefi.html>

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found at <http://genome.uiowa.edu/distribution/mousefi.html>

Seq primer: pYX-5.
 Location/Qualifiers
 1..710
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6848356"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH_BMAP_F10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCAGCAGC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 2.9%; Score 47; DB 14; Length 710;
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 826 CACTTCGAGTGACACTGATGACATTCAGCTCGAGATGGAACAGCA 872
 Db 407 CACTTCGAGTGACACTGATGACATTCAGCTCGAGATGGAACAGCA 453

RESULT 49
 CF728006
 LOCUS 716 bp mRNA linear EST 09-OCT-2003
 DEFINITION UI-M-HB0-ck1-n-24-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
 IMAGE:30550511 5', mRNA sequence.
 CF728006
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 716)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found at <http://genome.uiowa.edu/distribution/mousefi.html>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: PYX-5.
Location/Qualifiers
1. .716

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30550511"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP HB0"
/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."
ORIGIN

Query Match 2.9%; Score 47; DB 14; Length 716;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
|||||

DB 31 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 77
|||||

RESULT 50

BU0701958
LOCUS
DEFINITION UI-M-F10-bys-i-05-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone
IMAGE:5697868 5', mRNA sequence.
ACCESSION BU0701958
VERSION BU0701958.1 GI:23626281
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: PYX-5.
Location/Qualifiers
1. .725

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"

/clone="IMAGE:5697868"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_F10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 725;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
|||||

DB 373 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 419
|||||

RESULT 51

BU0770181
LOCUS
DEFINITION UI-M-F10-byt-1-02-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone
IMAGE:5702161 5', mRNA sequence.
ACCESSION BU0770181
VERSION BU0770181.1 GI:21978655
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: PYX-5
Location/Qualifiers
1. .727
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702161"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 727;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGCTCAGATGGAACAGCA 872
Db 454 CACTTCAGGTGACACTGATGACATTCAGCTCAGATGGAACAGCA 500

RESULT 52
LOCUS BQ769910 735 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-M-F10-byt-i-21-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
IMAGE:5702108 5', mRNA sequence.
ACCESSION BQ769910
VERSION BQ769910.1 GI:21978384
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 735)
NTH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1..735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702108"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="NIH BMAP F10"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES

source
1..735
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702108"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="NIH BMAP F10"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 735;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGCTCAGATGGAACAGCA 872
Db 386 CACTTCAGGTGACACTGATGACATTCAGCTCAGATGGAACAGCA 432

RESULT 53

LOCUS BQ769409 744 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-M-F10-byt-m-11-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
IMAGE:6401074 5', mRNA sequence.

ACCESSION BQ769409
VERSION BQ769409.1 GI:21977883
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 744)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1..744
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6401074"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="NIH BMAP F10"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES

source
1..744
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6401074"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="NIH BMAP F10"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.9%; Score 47; DB 13; Length 744;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;

```

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 872
      |||||||
DB 20 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 66

RESULT 54
CA327097 750 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FY0-ccy-g-01-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
DEFINITION IMAGE:6826106 5', mRNA sequence.
ACCESSION CA327097
VERSION CA327097.1 GI:24545195
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 750
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6826106"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/notes="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
to Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 2.9%; Score 47; DB 14; Length 750;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 872
      |||||||
DB 650 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 696

RESULT 55

```

```

CF735205 770 bp mRNA linear EST 10-OCT-2003
LOCUS UI-M-HB0-ckc-1-18-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
DEFINITION IMAGE:30615737 5', mRNA sequence.
ACCESSION CF735205
VERSION CF735205.1 GI:37631541
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 770
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30615737"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5,14.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/notes="Organ: Eye; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
to Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTATGTGAGT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN
Query Match 2.9%; Score 47; DB 14; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 872
      |||||||
DB 341 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 387

RESULT 56
CF728426 775 bp mRNA linear EST 09-OCT-2003
LOCUS UI-M-HB0-ckk-j-20-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
DEFINITION IMAGE:30547339 5', mRNA sequence.
ACCESSION CF728426
VERSION CF728426.1 GI:37602594
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 775)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 775
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30547339"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5,14.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/notes="Organ: Eye; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
to Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTATGTGAGT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN
Query Match 2.9%; Score 47; DB 14; Length 775;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 872
      |||||||
DB 341 CACTTCAGGTGACACTGAATGACATTCAGCTCAGATGGAACAGCA 387

RESULT 57

```


Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 775)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 775
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30547339"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5, 13.5, 14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP H80"
 /note="Organ: Eye; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the Not I site and the polyA tail
 sequence located between the Not I site and the polyA tail
 is TTATTCAGT. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System, supported by National
 Institute of Mental Health (NIMH)."

Query Match 2.9%; Score 47; DB 14; Length 775;
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
 Db 336 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 382

RESULT 57
 CB247992
 LOCUS
 DEFINITION UI-M-FIO-C-22-0-UI.r1 NIH BMAP_FIO Mus musculus CDNA clone
 IMAGE:6837167 5', mRNA sequence.

ACCESSION CB247992
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 779)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 779
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6837167"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP FIO"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the Not I site and the polyA tail
 sequence located between the Not I site and the polyA tail
 is CAGCAGCAGC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System, supported by National
 Institute of Mental Health (NIMH)."

Query Match 2.9%; Score 47; DB 14; Length 779;
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
 Db 475 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 521

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 779
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6837167"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP FIO"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the Not I site and the polyA tail
 sequence located between the Not I site and the polyA tail
 is CAGCAGCAGC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System, supported by National
 Institute of Mental Health (NIMH)."

Query Match 2.9%; Score 47; DB 14; Length 779;
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
 Db 475 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 521

RESULT 58
 CB328534
 LOCUS
 DEFINITION UI-M-FIO-cx-h-10-0-UI.r1 NIH BMAP_FIO Mus musculus CDNA clone
 IMAGE:6825755 5', mRNA sequence.

ACCESSION CB328534
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 798)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 798
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6825755"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP FIO"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the Not I site and the polyA tail
 sequence located between the Not I site and the polyA tail
 is CAGCAGCAGC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System, supported by National
 Institute of Mental Health (NIMH)."

Query Match 2.9%; Score 47; DB 14; Length 798;
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 872
 Db 475 CACTTCAGGTGACACTGATGACATTCAGTGCAGATGGAACAGCA 521

source

```

1. 798
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6825755"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F10"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

Query Match 2.9%; Score 47; DB 14; Length 798;

Best Local Similarity 100.0%; Pred. No. 5.5e-10; Indels 0; Gaps 0; Matches 47; Conservative 0; Mismatches 0;

QY 826 CACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 872

Db 430 CACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 476

RESULT 59

```

CA320886
LOCUS
DEFINITION
UI-M-FW0-ccb-h-16-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
IMAGE:6817313 5', mRNA sequence.
CA320886
VERSION
CA320886.1 GI:24538984
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 806)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
(This clone was contributed by the Brain Molecular Anatomy Project
(BMAP))
Seq primer: pYX-5.
Location/Qualifiers
1. 806
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6817313"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"

```

FEATURES

source

```

1. 806
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6817313"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"

```

```

/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F10"

```

```

/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

Query Match 2.9%; Score 47; DB 14; Length 806;

Best Local Similarity 100.0%; Pred. No. 5.5e-10; Indels 0; Gaps 0; Matches 47; Conservative 0; Mismatches 0;

QY 826 CACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 872

Db 465 CACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 511

RESULT 60

```

BF783468
LOCUS
DEFINITION
60211123BF1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4239513
5', mRNA sequence.
BF783468
ACCESSION
BF783468.1 GI:12088504
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 819)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9852 row: 1 column: 10
High quality sequence stop: 758.
Location/Qualifiers
1. 819
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4239513"
/clone_lib="NCI_CGAP_Kid14"

```

FEATURES

source

```

1. 819
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4239513"
/clone_lib="NCI_CGAP_Kid14"
/notes="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

```

ORIGIN

Query Match 2.9%; Score 47; DB 10; Length 819;

Best Local Similarity 100.0%; Pred. No. 5.5e-10; Indels 0; Gaps 0; Matches 47; Conservative 0; Mismatches 0;

```

826 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 872
|||||
608 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 654
|||||

RESULT 61
BI157950
LOCUS
DEFINITION
602920933F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5061033 5',
mRNA sequence.
ACCESSION
BI157950
VERSION
BI157950.1 GI:114617951
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 900)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1165 row: b column: 10
High quality sequence stop: 759.
FEATURES
Location/Qualifiers
1..900
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129_C57BL/6J_FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5061033"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam3"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sail;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

Query Match 2.9%; Score 47; DB 12; Length 900;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 872
|||||
Db 305 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 351
|||||

RESULT 62
CD469432
LOCUS
DEFINITION
CD469432 F09.gi A024 Stimulated peripheral blood leukocytes S2
Equus caballus cDNA clone LeukoS2_3_F09_A024 5', mRNA sequence.
ACCESSION
CD469432
VERSION
CD469432.1 GI:31390700
KEYWORDS
EST.
SOURCE
Equus caballus (horse)
ORGANISM
Equus caballus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

Query Match 2.9%; Score 47; DB 12; Length 900;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 826 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 872
|||||
Db 305 CACTTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCA 351
|||||

RESULT 62
CD469432
LOCUS
DEFINITION
CD469432 F09.gi A024 Stimulated peripheral blood leukocytes S2
Equus caballus cDNA clone LeukoS2_3_F09_A024 5', mRNA sequence.
ACCESSION
CD469432
VERSION
CD469432.1 GI:31390700
KEYWORDS
EST.
SOURCE
Equus caballus (horse)
ORGANISM
Equus caballus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

Query Match 2.8%; Score 46; DB 14; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1432 CCATATGACCTGAAACAAGAGGTACAGGACCTGAGTGTGTGGCC 1477
|||||
Db 283 CGCAATGACCTGAAACAAGAGGTACAGGACCTGAGTGTGTGGCC 328
|||||

RESULT 63
CA324134
LOCUS
DEFINITION
UI-M-FY0-CCO-F-03-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:5822244 5', mRNA sequence.
ACCESSION
CA324134
VERSION
CA324134.1 GI:24542232
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 778)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa

REFERENCE
1 (bases 1 to 754)
AUTHORS
Vandenplas, M., Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S.,
Moore, J.N., Liang, C., Sun, F., Sullivan, R., Shah, M. and Pratt, L.H.
An EST database from equine (Equus caballus) stimulated peripheral
blood leukocytes
JOURNAL
Unpublished (2003)
COMMENT
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; tissue and RNA were prepared in the Department of Large
Animal Medicine, University of Georgia; sequencing done in the
Laboratory for Genomics and Bioinformatics, University of Georgia.
Sequence ends have been trimmed to exclude vector and regions below
Phred quality 16. Three-prime sequences are presented as their
reverse complement and have been trimmed to exclude polyA.
Seq primer: Sug5 (CTTCGCTCTAAAGCTGGG).
Location/Qualifiers
1..754
/organism="Equus caballus"
/mol_type="mRNA"
/strain="thoroughbred"
/db_xref="taxon:9796"
/clone="LeukoS2_3_F09_A024"
/sex="male"
/tissue_type="blood"
/cell_type="leukocytes"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Stimulated peripheral blood leukocytes S2"
/notes="Organ: circulatory system; Vector: pME18S-FL3; from
Site 1: XhoI; Site 2: XhoI; The library was prepared from
polyA+ RNA from equine peripheral blood leukocytes
isolated from a healthy adult horse. The leukocytes were
stimulated for 4 hr with 10 ng/ml E. coli O55:B5 LPS.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5-prime
DraIII site is CACTGTGTG, 3-prime DraIII site is
CACCATGTG). XhoI excises the cDNA insert."

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (EMAP).

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
 1. 778
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6822244"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP FV0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; the library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGCAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 2.8%; Score 46; DB 14; Length 778;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 826 CACTTCAGGTGACACTGAATGACATTCAGTCAGTCAGTGAACAGC 871

Db 733 CACTTCAGGTGACACTGAATGACATTCAGTCAGTCAGTGAACAGC 778

RESULT 64

B0714219 942 bp mRNA linear EST 16-JUL-2002
 LOCUS AGENCOURT_8291711 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6308909
 5', mRNA sequence.

B0714219 GI:21853118

ACCESSION

B0714219 Mus musculus (house mouse)

KEYWORDS

SOURCE

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 942)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsof@mail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.

cDNA Library Preparation: ResGen, Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM13728 row: m column: 06

High quality sequence stop: 724.

FEATURES

source

Location/Qualifiers

1. 942
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6308909"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 129"
 /note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dT. Average insert size 2.2 kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
 Library."

ORIGIN

Query Match 2.7%; Score 45; DB 13; Length 942;
 Best Local Similarity 100.0%; Pred. No. 4.5e-09;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 828 CTTCCAGTGCACACTGAATGACATTCAGTCAGTCAGTGAACAGCA 872

Db 882 CTTCCAGTGCACACTGAATGACATTCAGTCAGTCAGTGAACAGCA 926

RESULT 65

B0533852

LOCUS

DEFINITION

B0533852 284 bp mRNA linear EST 25-JUN-2000
 duct Mus musculus cDNA clone 6720466K01.3', similar to L15344 Human
 high molecular weight B cell growth factor, mRNA sequence.

B0533852

EST. BB0533852.1 GI:8461000

KEYWORDS

SOURCE

ORGANISM

Mus musculus
 Mus musculus (house mouse)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 284)
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
 Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
 Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
 Okazaki, Y., Oho, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
 Shibata, K., Shibata, Y., Shigemoto, Y., Suzuki, H., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Sugahara, Y., Suzuki, H., Tsunoda, Y., Watahiki, A.,
 Takahashi, P., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A.,
 Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A.,
 Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

TITLE

JOURNAL

COMMENT

Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
 Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S.,
 Thermotabilization and thermotactivation of thermolabile enzymes by
 trchase and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
 Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.

/clone lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is TGAGAGACC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.5%; Score 41; DB 13; Length 723;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1174 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 1214
 Db 187 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 227

RESULT 68

CF741995 857 bp mRNA linear EST 10-OCT-2003
 LOCUS UI-M-HB0-clk-m-23-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
 DEFINITION IMAGE:30619990 5', mRNA sequence.

ACCESSION CF741995
 VERSION CF741995.1 GI:37638334
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 857)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
 1..857
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30619990"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_HB0"
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TTATGAAGT. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institutes of Mental Health (NIMH)."

ORIGIN

Query Match 2.5%; Score 41; DB 14; Length 857;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1174 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 1214
 Db 133 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 173

RESULT 69

BI102887 861 bp mRNA linear EST 26-JUN-2001
 LOCUS 602888485F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043572
 DEFINITION 5', mRNA sequence.

ACCESSION BI102887
 VERSION BI102887.1 GI:14553780
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 861)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1119 row: j Column: 21
 High quality sequence stop: 738.
 Location/Qualifiers
 1..861
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5043572"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Kid14"
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

FEATURES

Source
 1..861
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5043572"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Kid14"
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 2.5%; Score 41; DB 12; Length 861;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1174 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 1214
 Db 266 GAGAAGTTTGAGGAGTTCAGACACACACTTCCAAAAGCAG 306

RESULT 70

BUI51934 954 bp mRNA linear EST 03-SEP-2002
 LOCUS AGENCOURT_8753209 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6394355
 DEFINITION 5', mRNA sequence.

ACCESSION BUI151934
 VERSION BUI151934.1 GI:22665466
 KEYWORDS ESI.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 954)
 AUTHORS NIH-MGC <http://mgi.nsl.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 CDNA Library Preparation: ResGen, Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM13887 row: e column: 12
 High quality sequence stop: 512.
 Location/Qualifiers
 1..954
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6394355"
 /clone_lib="NIH MGC 130"
 /notes="Organ: otcysts; Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 2.5%; Score 41; DB 13; Length 954;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1174 GAGAGTTTGGAGGTTCCAGACACACTTCCAAAGCAG 1214
 Db 304 GAGAGTTTGGAGGTTCCAGACACACTTCCAAAGCAG 344
 RESULT 71
 BB052776
 LOCUS
 DEFINITION
 BB052776 RIKEN full-length enriched, 12 days embryo male wolffian duct Mus musculus cDNA clone 6720456M09 3' similar to L15344 Human high molecular weight B cell growth factor, mRNA sequence.
 BB052776
 VERSION BB052776.1 GI:8459924
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 210)
 AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Iori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sugabe, Y., Sugahara, Y., Suzuki, H., Shinagawa, A., Shiraki, T., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)
 Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Tomaru, Y., Carninci, P., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES source

Location/Qualifiers
 1..210
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="6720456M09"
 /sex="male"
 /tissue_type="wolffian duct includes surrounding region"
 /dev_stage="12 days embryo"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 12 days embryo male wolffian duct"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Query Match 2.4%; Score 40; DB 10; Length 210;
 Best Local Similarity 100.0%; Pred. No. 7.3e-07;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1174 GAGAGTTTGGAGGTTCCAGACACACTTCCAAAGCA 1213
 Db 138 GAGAGTTTGGAGGTTCCAGACACACTTCCAAAGCA 177
 RESULT 72
 AT1768249 414 bp mRNA linear EST 20-DEC-1999
 LOCUS W938a04.x1 Soares NSF_P8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:237162 3' similar to SW:IL14 HUMAN P40222 INTERLEUKIN-14 PRECURSOR ; contains element PTR7 PTR7 repetitive element ; mRNA sequence.

ACCESSION AI768249
 VERSION AI768249.1 GI:5234758
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 414)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 640 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 121.
 FEATURES
 Location/Qualifiers
 1..414
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2371662"
 /lab_host="DH10B"
 /clone_lib="Soares NSP_F8_9W_OT_PA_P_S1"
 /notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NB2HP pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HPF-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."
 ORIGIN
 Query Match 2.4%; Score 40; DB 9; Length 414;
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1155 ACACCTTGCCTATACACAGAGAGTTTGAGGAGTTCCAG 1194
 Db 414 ACACCTTGCCTATACACAGAGAGTTTGAGGAGTTCCAG 375
 RESULT 73
 CB244426
 LOCUS
 DEFINITION UT-M-FYO-cdg-i-09-0-UI.r1 NIH_BMAP_FYO Mus musculus CDNA clone IMAGE:6833074 5', mRNA sequence.
 ACCESSION CB244426
 VERSION CB244426.1 GI:28366070
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 727)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP).
 Seq primer: pYX-5.
 Location/Qualifiers
 1..727
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6833074"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FYO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."
 ORIGIN
 Query Match 2.4%; Score 39; DB 14; Length 727;
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 834 GGTGACATGATGACATTCAGTTCAGTGCAGTGCAGACGCA 872
 Db 513 GGTGACATGATGACATTCAGTTCAGTGCAGTGCAGACGCA 551
 RESULT 74
 BG062308
 LOCUS
 DEFINITION L0950F05-5 NIA Mouse Newborn Kidney CDNA Library2 (Short) Mus musculus CDNA clone L0950F05 5', mRNA sequence.
 ACCESSION BG062308
 VERSION BG062308.2 GI:31576471
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 409)
 AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
 JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
 MEDLINE 21429098
 PUBMED 11541199
 COMMENT On Jan 25, 2001 this sequence version replaced gi:12532655.
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 niaEST (http://igsun.grc.nia.nih.gov/cdna/cdna.html)

Plate: L0950 row: F column: 05
 Seq primer: -21M13 Reverse
 High quality sequence stop: 409
 POLYA=No.

FEATURES

Location/Qualifiers
 1..409
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:L0950F05-5"
 /db_xref="taxon:10090"
 /clone="L0950F05"
 /tissue_type="Newborn Kidney"
 /dev_stage="Newborn"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library2
 (Short)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a short-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). In brief, double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-PGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from 26 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-L. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.5 kb. The library was constructed by Yulan Piao(NIA)."

ORIGIN

Query Match 2.3%; Score 38; DB 10; Length 409;
 Best Local Similarity 100.0%; Pred. No. 6.3e-06;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1258 AAGAAGATCAAGAGCTGGAGAAAGAACACCACCATGTA 1295
 DB 169 AAGAAGATCAAGAGCTGGAGAAAGAACACCACCATGTA 206

RESULT 75

CA885982
 LOCUS B012A04-5N NIA Mouse Neural Stem Cell EST 20-DEC-2002
 DEFINITION Library (Long) Mus musculus cDNA clone NIA:B012A04 IMAGE:30096099
 5', mRNA sequence.

ACCESSION CA885982
 VERSION CA885982.1 GI:273337531
 KEYWORDS EST.

SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 608)
 AUTHORS Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Alba,K., Vescovi,A.L. and Ko,M.S.H.

TITLE Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated) cDNA Library (Long)

JOURNAL Unpublished (2002)
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0124 row: A column: 04
 Seq primer: -21M13 Reverse
 High quality sequence stop: 608
 POLYA=No.

FEATURES

Location/Qualifiers
 1..608
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CD1"
 /db_xref="niaEST:B0124A04-5N"
 /db_xref="taxon:10090"
 /clone="NIA:B0124A04 IMAGE:30096099"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Neural Stem Cell (Differentiated)
 cDNA Library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-PGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.2 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 2.3%; Score 38; DB 14; Length 608;
 Best Local Similarity 100.0%; Pred. No. 6.5e-06;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1258 AAGAAGATCAAGAGCTGGAGAAAGAACACCACCATGTA 1295
 DB 3 AAGAAGATCAAGAGCTGGAGAAAGAACACCACCATGTA 40

RESULT 76

CF898484
 LOCUS A0240A06-5 NTA Mouse Embryonic Germ Cell cDNA Library EST 04-NOV-2003
 DEFINITION subtrated Mus musculus cDNA clone NIA:A0240A06 IMAGE:30732197 5', mRNA sequence.

ACCESSION CF898484
 VERSION CF898484.1 GI:38165533
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 675)

AUTHORS Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.

TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE 21429098

PUBMED 11544199

COMMENT Contact: Dawood B. Dudekula

Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdaelg@nigam.nih.gov
Plate: A0240 row: A column: 06
Seq primer: M13 Reverse
High quality sequence stop: 675
POLYA=No.

FEATURES

source

1. 675
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57B16"
/db_xref="nia:EST:A0240A06-5"
/db_xref="taxon:10090"
/clone="NIA:A0240A06 IMAGE:30732197"
/sex="male"
/dev_stage="embryonic day 8"
/lab_host="DH10B"
/clone_lib="NIA Mouse Embryonic Germ Cell cDNA Library (Long_subtracted)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). EG cells were obtained from Dr. Brigid L.M. Hogan and RNA was prepared by Dr. Mark G. Carter (NIH/NIA-IRP). EG cells were cultured at 37. C, 5% CO2 in DMEM supplemented with 15% ES cell-qualified FBS, 0.1mM non-essential amino acids, 2 mM glutamine, penicillin/streptomycin, 1 mM sodium pyruvate, 0.1 mM beta-mercaptoethanol, and 10⁷ units of LIF per liter. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCCGAGCGCCGCCCTTTTCTTTT-3'] from 2.5 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LI-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were double digested with NotI and SalI enzymes, then purified by phenol/chloroform and Centricon 100. The cDNA mixture was subjected to a special subtraction procedure by Dr. Kazuhiro Kondo at Aisin Cosmos. Then the subtracted cDNAs were cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2kb. The library was constructed by Yulan Piao and Kazuhiro Kondo."

ORIGIN

Query Match 2.3%; Score 38; DB 14; Length 675;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1258 AAGAAGATCAAGAGCTGGAGAAAGAAACCAACCATGTA 1295
Db 284 AAGAAGATCAAGAGCTGGAGAAAGAAACCAACCATGTA 321

RESULT 77
BG864439 821 bp mRNA linear EST 29-MAY-2001
LOCUS 602798647F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4919626 5',
DEFINITION mRNA sequence.
ACCESSION BG864439
VERSION BG864439.1 GI:14214977
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 821)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10834 row: n column: 11
High quality sequence start: 2
High quality sequence stop: 709.

FEATURES

source

1. 821
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4919626"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

ORIGIN

Query Match 2.3%; Score 38; DB 12; Length 821;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1258 AAGAAGATCAAGAGCTGGAGAAAGAAACCAACCATGTA 1295
Db 68 AAGAAGATCAAGAGCTGGAGAAAGAAACCAACCATGTA 105

RESULT 78

BG601883

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 466)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrner, S.C.,

Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,

Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F.,

Quackenbush, J., and Keele, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

11282978

PUBLISHED

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 42 row: L column: 11
 Seq primer: ATTAGTGACACTATAG.

FEATURES

Location/Qualifiers
 1..466

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 3BOV"

/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."

ORIGIN

Query Match 2.3%; Score 37; DB 10; Length 466;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 GCTGAACCCGAGATGCGAGAGAGTCCCGACCTATG 349

Db 409 GCTGAACCCGAGATGCGAGAGAGTCCCGACCTATG 445

RESULT 79

CF167914

LOCUS

DEFINITION B0791G03-5 NIA Mouse Embryonic Germ Cell cDNA Library (Long) Mus

musculus cDNA clone NIA:B0791G03 IMAGE:30466634 5', mRNA sequence.

ACCESSION CF167914

VERSION CF167914.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 497)

Piao,Y., Ko,N.T., Lim,M.X. and Ko,M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

21429098

11544199

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: B0791 row: G column: 03

Seq primer: M13 Reverse

High quality sequence stop: 497

POLVA=NO.

FEATURES

source

1..497

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="niaEST:B0791G03-5"

/db_xref="taxon:10090"

/clone="NIA:B0791G03 IMAGE:30466634"

/sex="male"

/dev_stage="embryonic day 8"

/lab_host="DH10B"
 /clone_lib="NIA Mouse Embryonic Germ Cell cDNA Library
 (Long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
 RNAs were obtained from Dr Mark G. Carter (NIH/NIA-IRP).
 EG cells were cultured at 37° C, 5% CO2 in DMEM
 supplemented with 15% ES cell-qualified FBS, 0.1mM
 non-essential amino acids, 2 mM glutamine,
 penicillin/streptomycin, 1 mM sodium pyruvate, 0.1 mM
 beta-mercaptoethanol, and 1000000 units of LIF per liter.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen].
 5'-PGACTAGTCTAGATCGGAGCGCCCTTTTCTTTT-3' from
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lona-linker L1-SalI, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer SalI-8. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 4.0 kb. The library was
 constructed by Yulan Piao."

ORIGIN

Query Match 2.2%; Score 36; DB 14; Length 497;

Best Local Similarity 100.0%; Pred. No. 5.2e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 CAAGATCAAGAGCTGGAGAGAAACACCATGTA 1295

Db 1 GAAGATCAAGAGCTGGAGAGAAACACCATGTA 36

RESULT 80

BU613265

LOCUS

DEFINITION

UI-M-EMO-cay-n-23-0-UI-r1 NIH EMAP_EMO Mus musculus cDNA clone

UI-M-EMO-cay-n-23-0-UI 5', mRNA sequence.

ACCESSION BU613265

VERSION BU613265.1

KEYWORDS GI:23279480

SOURCE EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 725)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgsabp-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bentosoares@iowa.edu

This clone was contributed by the Brain Molecular Anatomy Project

(BNAP)

Seq primer: pYX-5.

Location/Qualifiers

1..725

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UI-M-EW0-cay-n-23-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP EW0"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 2.1%; Score 34; DB 13; Length 725;
 Best Local Similarity 100.0%; Pred. NO. 0.00043;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AGATCAGAGACCTGGAGAAAGAACACCATGTA 1295
 |||||
 Db 1 AGATCAGAGACCTGGAGAAAGAACACCATGTA 34

RESULT 81

AZ391393 640 bp DNA linear GSS 03-OCT-2000
 LOCUS M0153A03R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
 DEFINITION clone UGCGIM0153A03 R, genomic survey sequence.

ACCESSION AZ391393
 VERSION AZ391393.1 GI:10506436
 KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhauser, A., and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0153 row: A column: 03

Seq primer: CACACGGAACACTATGACC

Class: plasmid ends

High quality sequence stop: 640.

Location/Qualifiers

1. 640

FEATURES

source

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0153A03"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UGCGIM library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 [gi4732114|gb|AF129072.1], a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 2.0%; Score 32; DB 28; Length 640;
 Best Local Similarity 100.0%; Pred. NO. 0.0035;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1183 GAGGAGTTCACAGACACATTTCCAAAAGCAG 1214
 |||||
 Db 558 GAGGAGTTCACAGACACATTTCCAAAAGCAG 589

RESULT 82

CB171858/c

LOCUS

DEFINITION

CB171858

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 745)

Adelson, D.L. and Gill, C.A.

Bovine ESTs (Adelson and Gill)

Unpublished (2003)

JOURNAL

COMMENT

Contact: David L. Adelson

Animal Breeding and Genetics

Texas A&M University

Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,

USA

Tel: 9798452616

Fax: 9798456970

Email: david.adelson@tamu.edu.

Location/Qualifiers

1. 745

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="thyroid"

/clone_lib="CSEQFN46 thyroid and parathyroid"

/note="Organ: thyroid and parathyroid; Vector: pBluescript

SK+; Site:1: NotI; Site:2: EcoRI; sequence 5' of the

insert (5'-NNN...NNNinsert)

CGCAATGAGCTCCCGGCGCGCGCGCGCTCGAG. Sequence 3' of

the inserts (AAGATTCGATCATGATTCGATACCGTCCGACCTCGAG.

normalized Rd 2 library, sequenced 3' with M13R primer."

ORIGIN

```

Query Match      2.0%; Score 32; DB 14; Length 745;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 AAGGAGATCACGTTGCTGATGCAGACATTGAA 539
    |||||
Db 540 AAGGAGATCACGTTGCTGATGCAGACATTGAA 509

RESULT 83
BX874904          512 bp mRNA linear EST 17-DEC-2003
LOCUS             BX874904
DEFINITION        OX874904 AGENAE Rainbow trout multi-tissues-normalized (tcbk)
VERSION           BX874904
KEYWORDS          OX874904
SOURCE            BX874904.1 GI:40003449
ORGANISM          Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 512)
AUTHORS Govoroun,M., Guiguen,Y. and Le Gac,F.
TITLE Construction and primary characterization of normalized cDNA
JOURNAL libraries in rainbow trout, Oncorhynchus mykiss
COMMENT Unpublished (2003)
Contact: Guiguen Y
INRA - SCRIBE
Campus de beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: fann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0006 row: 9 column: 17
Seq primer: M13R.
FEATURES
    source
    1..512
        /organism="Oncorhynchus mykiss"
        /mol_type="mRNA"
        /db_xref="taxon:8022"
        /clone="tcbk0006c.g.17"
        /tissue_type="multi-tissues"
        /dev_stage="from embryos to adults"
        /lab_host="DH10B"
        /clone_lib="AGENAE Rainbow trout multi-tissues-normalized
        (tcbk)"
        /notes="Vector: pT73D-pac; Rainbow trout
        multi-tissues-normalized + 2 subtractions; Clone
        distribution : AGENAE Resource centre, Francois PIUMI,
        Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
        Etude du genome (LREG), Domaine de Vilvert, 78352,
        Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
        (0) 1.34.65.22.73"

Query Match      1.8%; Score 30; DB 13; Length 512;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1166 TATACACAGAGAGTTTGAGGAGTTCCAGA 1195
    |||||
Db 235 TATACACAGAGAGTTTGAGGAGTTCCAGA 264

RESULT 84
CA355951          619 bp mRNA linear EST 05-NOV-2002
LOCUS             CA355951
DEFINITION        627950 NCCOWA 1RT Oncorhynchus mykiss cDNA clone IRT8M06_C_G03 5',
                    mRNA sequence.
ACCESSION         CA355951

```

```

CA355951.1 GI:24601138
EST.
Oncorhynchus mykiss (rainbow trout)
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 619)
AUTHORS Rexroad,C.E. and Keele,J.W.
TITLE Sequence analysis of a rainbow trout normalized cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@cccw.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified by
cross_match v0.990329.
Seq primer: ACCGATACAAATTCACACAGGA.
FEATURES
    source
    1..619
        /organism="Oncorhynchus mykiss"
        /mol_type="mRNA"
        /db_xref="taxon:8022"
        /clone="IRT8M06_C_G03"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /clone_lib="NCCOWA 1RT"
        /note="Vector: PCWV SPORT6; Site 1: Not1; Site 2: Sal1;
        Library made from pooled tissue from brain, gill, liver,
        spleen, muscle, and kidney."

Query Match      1.8%; Score 30; DB 14; Length 619;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1166 TATACACAGAGAGTTTGAGGAGTTCCAGA 1195
    |||||
Db 258 TATACACAGAGAGTTTGAGGAGTTCCAGA 287

RESULT 85
AI877686          761 bp mRNA linear EST 07-JUN-2001
LOCUS             AI877686
DEFINITION        f50d09.y1 Zebrafish Washu MFPMG EST Danio rerio cDNA clone
                    IMAGE:3724817 5', similar to SW:1114_HUMAN P40222 INTERLEUKIN-14
                    PRECURSOR, mRNA sequence.
ACCESSION         AI877686
VERSION           AI877686.1 GI:5551735
KEYWORDS          EST.
SOURCE            Danio rerio (zebrafish)
ORGANISM          Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 761)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
                    Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
                    Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
                    Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
                    Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
                    Waterston,R. and Wilson,R.
                    Washu Zebrafish EST Project 1998
                    Unpublished (1998)
                    Other ESTs: f50d09.xl
                    Contact: Stephen L. Johnson
                    Washington University School of Medicine
                    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                    Tel: 314 286 1800
                    Fax: 314 286 1810

```

Email: zbrafish@watson.wustl.edu
 cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
 Matthew Clark. DNA Sequencing by: Washington University Genome
 Sequencing Center Clone Distribution: Genome Systems, St. Louis,
 Missouri (web address: www.genomesystems.com) (email contact:
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama
 (web address: www.resgen.com) (email contact: info@resgen.com) and
 Ressourcenzentrum Primatendarbank, Berlin, Germany (web address:
 www.rzpd.de)

Possible reversed clone: similarity on wrong strand

Seq primer: T3 Et from Amersham

High quality sequence stop: 457.

FEATURES

source

Location/Qualifiers
 1..761
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:3724817"
 /sex="mixed"
 /tissue types="26 somite embryos, adult livers, shield
 stage embryos"
 /lab_host="X11-blue MRP"
 /clone_lib="Zebrafish WashU MPIMG EST"
 /note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; 1st
 strand cDNA was primed with a Not I - oligo (dT)15 primer
 (5'PGACTAGTTCTAGATCGGAGCGCCCTTTTCTTTT3');
 double-stranded cDNA was ligated to Sal I adaptors (BRL),
 digested with Not I and cloned into the Not I and Sal I
 sites of the pSPOR1 vector (BRL). Library was constructed
 by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
 Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
 analysis were selected following oligonucleotide
 hybridization fingerprinting of arrayed clones from
 zebrafish late somitogenesis (26 ss), adult liver or
 embryonic shield stage (5.6 h) libraries. Fingerprint
 data were used to computationally cluster cDNAs, and a
 single cDNA from each cluster was chosen for sequencing.
 In some cases multiple members of the same cluster were
 sequenced to assess clustering parameters or single clones
 were sequenced additional times to assess quality
 control."

ORIGIN

Query Match 1.8%; Score 30; DB 9; Length 761;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 888 CAAGCTCGCCCAAGAGAACATGGAGTGGC 917

Db 681 CAAGCTCGCCCAAGAGAACATGGAGTGGC 710

RESULT 86

LOCUS BX855901

DEFINITION BX855901 AGENAE Rainbow trout multi-tissues substracted library
 (tcay) Oncorhynchus mykiss cDNA clone tcay0039b.j.10 Sprin, mRNA
 sequence.

ACCESSION BX855901

VERSION BX855901.1 GI:39952895

KEYWORDS EST.

SOURCE

ORGANISM

Oncorhynchus mykiss (rainbow trout)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (Bases 1 to 780)

Govoroun, M., Guiguen, Y. and Le Gac, F.

Construction and primary characterization of normalized cDNA

libraries in rainbow trout, Oncorhynchus mykiss

Unpublished (2003)

Contact: Guiguen Y

INRA - SCRIBE

Campus de beaulieu, RENNES cedex, 35042, France

Tel: 02.23.48.50.09

Fax: 02.23.48.50.20

Email: Yann.Guiguen@beaulieu.rennes.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us

at signenasupport@jouy.inra.fr to obtain the chromatogram of this

sequence.

Plate: 0039 row: j column: 10

Seq primer: M13R.

FEATURES

source

Location/Qualifiers

1..780

/organism="Oncorhynchus mykiss"

/mol_type="mRNA"

/db_xref="taxon:8022"

/clone="tcay0039b.j.10"

/tissue types="adipose tissue, blood, brain,

differentiating gonads, gills, interrenal, intestine,

kidney, liver, muscle, ovary, pituitary, testis"

/dev stage="from embryos to adults"

/lab host="DH10B"

/clone lib="AGENAE Rainbow trout multi-tissues substracted

library (tcay)"

/note="Vector: pT7T3D-pac; Rainbow trout multi-tissues -

normalized + 1 subtraction (tcay): Clone distribution :

AGENAE Resource centre. Francois PIUMI,

Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et

Etude du genome (LREG), Domaine de Vilvert, 78352,

Jouy-en-Josas cedex, FRANCE"

Query Match 1.8%; Score 30; DB 13; Length 780;

Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1166 TATACACAGAGAGCTTTGAGGAGTCCAGA 1195

Db 664 TATACACAGAGAGCTTTGAGGAGTCCAGA 693

RESULT 87

LOCUS BF783480

DEFINITION BF783480

5' mRNA sequence.

ACCESSION BF783480

VERSION BF783480.1 GI:12088516

KEYWORDS EST.

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (Bases 1 to 464)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LAM9852 row: 1 column: 14

High quality sequence stop: 415.

Location/Qualifiers

1..464

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4239517"

FEATURES

source

/lab host="DH10B (T1 phage-resistant)"
 /clone lib="NCI_CGAP_Kid14"
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library. |"

ORIGIN

Query Match 1.8%; Score 29; DB 10; Length 464;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
 Db 330 CTGAGTACCCAGAGGAGAGCTGGCTGC 358

RESULT 88
 CD564467
 LOCUS
 DEFINITION B0479D07-5 NIA Mouse E6.5 Whole Embryo cDNA Library (Long) Mus EST 11-JUN-2003
 musculus cDNA clone NIA:B0479D07 IMAGE:30446250 5', mRNA sequence.
 CD564467
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 503)
 TITLE Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
 MEDLINE 21429098
 PUBMED 11544199
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cna@nigun.grc.nia.nih.gov
 Plate: B0479 row: D column: 07
 Seq primer: M13 Reverse
 High quality sequence stop: 503
 POLYA-No.

FEATURES
 source Location/Qualifiers
 1..503
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:B0479D07-5"
 /db_xref="taxon:10090"
 /clone="NIA:B0479D07 IMAGE:30446250"
 /tissue_type="E6.5 whole embryo"
 /dev_stage="whole embryo including extraembryonic tissues
 at 6.5-days postcoitum"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse E6.5 Whole Embryo cDNA Library
 (Long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site: 1: SalI;
 Site: 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://nigun.grc.nia.nih.gov/cDNA).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001) [PMID: 11544199]). Total
 RNAs were extracted from a pool of 7 embryos at 6.5-days
 postcoitum. Double-stranded cDNAs were synthesized with an
 Oligo(dT) primer.
 5'-PGACTAGTCTAGATCGAGCGCGCCCTTTT-3'
 from 0.53 ug of total RNA, treated with T4 DNA polymerase,
 and purified by ethanol-precipitation. The cDNAs were
 ligated to lone-linker LL-SalI, purified by

phenol/chloroform, and separated from free linkers by
 Centricon 100. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-S. The products were purified
 by phenol/chloroform and Centricon 100. The cDNAs were
 digested with SalI and NotI enzymes and cloned into
 SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E.
 coli host was transformed with the ligation mixture by the
 standard chemical method. The average insert size is about
 2.3kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 1.8%; Score 29; DB 14; Length 503;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
 Db 239 CTGAGTACCCAGAGGAGAGCTGGCTGC 267

RESULT 89
 CC249288
 LOCUS
 DEFINITION RR091 BayGenomics Gene Trap Library pGT2Lxf Mus musculus cDNA,
 mRNA sequence.
 CC249288
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 571)
 TITLE BayGenomics.
 JOURNAL http://baygenomics.ucsf.edu/
 COMMENT Unpublished (2001)
 Contact: BayGenomics
 Bay Area Functional Genomics Consortium (BayGenomics)
 Email: info@baygenomics.ucsf.edu
 Sequence tag generated by 5' RACE of total RNA from gene trap ES
 cell line. ES cell lines harboring insertion mutation of target
 gene are available upon request from BayGenomics. Annotation
 information available from http://baygenomics.ucsf.edu/cgi-bin/
 BaySearch.py?Option=EXACT&TYPE=CELL_LINE&KEY=RRF091
 Class: Gene Trap.

FEATURES
 source Location/Qualifiers
 1..571
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129 OLA"
 /db_xref="taxon:10090"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /clone_lib="BayGenomics Gene Trap Library pGT2Lxf"
 /note="Vector: pGT2Lxf"

ORIGIN

Query Match 1.8%; Score 29; DB 28; Length 571;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
 Db 349 CTGAGTACCCAGAGGAGAGCTGGCTGC 377

RESULT 90
 BI065339
 LOCUS
 DEFINITION pgfin.pk004.e24 normalized chicken fat cDNA library Gallus gallus
 cDNA clone pgfin.pk004.e24 5', similar to emb|CAB75615.1 (AL049795)
 dJ622L5.2 (novel protein) [Homo sapiens]G, mRNA sequence.

```

ACCESSION      BI065339
VERSION        BI065339.1  GI:14472861
KEYWORDS       EST.
SOURCE         Gallus gallus (chicken)
ORGANISM       Gallus gallus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.
REFERENCE      1 (bases 1 to 594)
AUTHORS        Cogburn, D.A., Morgan, R.W. and Burnside, J.
TITLE          Chicken ESTs from fat
JOURNAL        Unpublished (2001)
COMMENT        Contact: Larry A. Cogburn
                University of Delaware
                Townsend Hall, Newark, DE 19717, USA
                Tel: 302-831-1335
                Fax: 302-831-2822
                Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES       source
                1..594
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /db_xref="taxon:9031"
                /clone="pgf1n.pk004.e24"
                /sex="Male and Female"
                /tissue types="fat"
                /lab host="E.coli EMDH103"
                /clone lib="normalized chicken fat cDNA library"
                /note="Vector: pSPORT1"

ORIGIN
Query Match      1.8%; Score 29; DB 12; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1108 CAGAGGATGTGTGAGCTGATGAGCAGCA 1136
        |||||
Db       478 CAGAGGATGTGTGAGCTGATGAGCAGCA 506

RESULT 91
LOCUS      CD550956
DEFINITION B0321E01-5 NIA Mouse B9.5 Whole Embryo cDNA Library (Long) Mus
            musculus cDNA clone NIA:B0321E01 IMAGE:30431088 5', mRNA sequence.
ACCESSION  CD550956
VERSION     CD550956.1  GI:31598687
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 618)
AUTHORS     Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE       Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
JOURNAL     Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE     21429098
PUBMED      11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@gsun.grc.nia.nih.gov
            Plate: B0321 row: E column: 01
            Seq primer: M13 Reverse
            High quality sequence stop: 618
            POLYA-No.

FEATURES     source
                Location/Qualifiers
                1..618
                /organism="Mus musculus"
                /mol_type="mRNA"

```

```

/strain="C57BL/6J"
/db_xref="niaEST:B0321E01-5"
/db_xref="taxon:10090"
/clone="NIA:B0321E01 IMAGE:30431088"
/tissue types="B9.5 whole embryo"
/dev stage="whole embryo including extraembryonic tissues
at 9.5-days postcoitum"
/lab host="DH10B"
/clone lib="NIA Mouse B9.5 Whole Embryo cDNA Library
(Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001)). (PMID: 11544199). Total
RNAs were extracted from a pool of 16 embryos at 9.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer (Invitrogen):
5'-PGACTAGTTCGATCGAGCGCGCCCTTTTCTTTT-3'
and purified by ethanol-precipitation. The cDNAs were
ligated to lone-linker LI-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.0Kb. The library was constructed by Yulan Piao."

ORIGIN
Query Match      1.8%; Score 29; DB 14; Length 618;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      544 CTGAGTACCCAGAGAGAGCTGGCTGC 572
        |||||
Db       566 CTGAGTACCCAGAGAGAGCTGGCTGC 594

RESULT 92
LOCUS      CD542146
DEFINITION B0238A11-5 NIA Mouse Embryonic Germ Cell cDNA Library (Long) Mus
            musculus cDNA clone NIA:B0238A11 IMAGE:30107050 5', mRNA sequence.
ACCESSION  CD542146
VERSION     CD542146.1  GI:31589881
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 629)
AUTHORS     Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE       Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
JOURNAL     Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE     21429098
PUBMED      11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@gsun.grc.nia.nih.gov
            Plate: B0238 row: A column: 11
            Seq primer: M13 Reverse
            High quality sequence stop: 629
            POLYA-No.

```


FEATURES

source
 1. .629
 /Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:B0238A11-5"
 /db_xref="taxon:10090"
 /clone="NIA:B0238A11 IMAGE:30107050"
 /sex="male"
 /dev_stage="embryonic day 8"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Embryonic Germ Cell cDNA Library (Long)"
 /notes="vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Mark G. Carter (NIH/NIA-IRP). EG cells were cultured at 37. C, 5% CO2 in DMEM supplemented with 15% ES cell-qualified FBS, 0.1mM penicillin/streptomycin, 1 mM sodium pyruvate, 0.1 mM beta-mercaptoethanol, and 10000000 units of LIF per liter. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
 5'-pGACTGTTTATGATCGGAGCGCGCCCTTTT-3' from 2.5 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 4.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 1.8%; Score 29; DB 14; Length 629;
 Best Local Similarity 100.0%; Pred. No. 0.079;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 544 CTGATGACCCAGAGGAGAGCTGGCTGC 572

Db 598 CTGATGACCCAGAGGAGAGCTGGCTGC 626

RESULT 93

CE687555/c
 LOCUS
 DEFINITION
 tigr-gss-dog-17000314567969 Dog Library Canis familiaris genomic,
 genomic survey sequence.

ACCESSION CE687555

VERSION CE687555.1 GI:37006625

KEYWORDS GSS.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 653)

REFERENCE

AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,

Rusch, D.B., DeCher, A.L., Pop, M., Wang, W., Fraser, C.M. and

Venter, J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

MEDLINE 22875432

PUBMED 14512627

COMMENT Contact: Kirkness EF

The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.

FEATURES

source
 1. .653
 /Location/Qualifiers
 /organism="Canis familiaris"
 /mol_type="Genomic DNA"
 /strains="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 1.8%; Score 29; DB 29; Length 653;
 Best Local Similarity 100.0%; Pred. No. 0.079;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 205 COTGATGTCCTCTGAGGAGCTGAGCGGCCA 233

Db 353 COTGATGTCCTCTGAGGAGCTGAGCGGCCA 325

RESULT 94

BY749082

LOCUS

DEFINITION

BY749082 661 bp mRNA linear EST 17-DEC-2002
 BY749082 RIKEN full-length enriched, NON-derived cDNA clone F630224A22 5', mRNA
 dendritic cells Mus musculus

ACCESSION BY749082

VERSION BY749082.1 GI:27178283

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 661)

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Ictani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC Building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source
1. .661
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F630224A22"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched, NOD-derived CD11c
+ve dendritic cells"

ORIGIN

Query Match 1.8%; Score 29; DB 13; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
|||||
DB 621 CTGAGTACCCAGAGGAGAGCTGGCTGC 649
|||||

RESULT 95

BU701946 669 bp mRNA linear EST 15-JUL-2003
LOCUS UI-M-F10-bys-g-01-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
DEFINITION IMAGE:5697816 5', mRNA sequence.
BU701946
ACCESSION BU701946.1 GI:23626256
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 669)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source

Seq primer: pyX-5.
Location/Qualifiers
1. .669
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5697816"
/tissue_type="whole brain"
/dev_stages="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F10"
/note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
is located between the Not I site and the polyA tail
is CAGCCAGCAGC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 1.8%; Score 29; DB 13; Length 669;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
|||||
DB 594 CTGAGTACCCAGAGGAGAGCTGGCTGC 622
|||||

RESULT 96

CD578619 674 bp mRNA linear EST 09-JUL-2003.
LOCUS UI-M-FY0-cft-i-16-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
DEFINITION IMAGE:6852281 5', mRNA sequence.
CD578619
ACCESSION CD578619.1 GI:31743010
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 674)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5,
 Location/Qualifiers
 1. .674

FEATURES

source

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I; Site:2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."
```

ORIGIN

```
Query Match 1.8%; Score 29; DB 14; Length 674;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
DB 549 CTGAGTACCCAGAGGAGAGCTGGCTGC 577
```

RESULT 97

```
BU611554
LOCUS BU611554 680 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-M-F10-cav-o-06-0-UI.r1 NIH BMAP F10 Mus musculus cDNA clone
ACCESSION BU611554
VERSION BU611554.1 GI:23277769
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 680)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
```

FEATURES

source

```
Location/Qualifiers
1. .680
/organism="Mus musculus"
```

```
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-F10-cav-o-06-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I; Site:2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCAGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

```
Query Match 1.8%; Score 29; DB 13; Length 680;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGGAGAGCTGGCTGC 572
DB 615 CTGAGTACCCAGAGGAGAGCTGGCTGC 643
```

RESULT 98

```
BU054710
LOCUS BU054710 686 bp mRNA linear EST 26-AUG-2002
DEFINITION UI-M-FD0-bzi-k-16-0-UI.r1 NIH BMAP_FD0 Mus musculus cDNA clone
ACCESSION BU054710
VERSION BU054710.1 GI:22494787
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 686)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
```

FEATURES

source

```
Location/Qualifiers
1. .686
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:640103"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FD0"
```

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGACC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.8%; Score 29; DB 13; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGAGAGCTGGCTGC 572
|||||
Db 536 CTGAGTACCCAGAGAGAGCTGGCTGC 564
|||||

RESULT 99

CF535545
LOCUS
DEFINITION UI-M-GHO-cgr-g-12-0-UI.r1 NIH BMAP_GHO Mus musculus cDNA clone
IMAGE:30534587 5', mRNA sequence.
ACCESSION CF535545
VERSION
KEYWORDS
SOURCE EST. GI:34587513
Mus musculus (house mouse)
ORGANISM
REFERENCE 1 (bases 1 to 686)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

source
Seq primer: pYX-5.
Location/Qualifiers
1..686
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30534587"
/tissue_type="Whole brain"
/dev_stages="1, 5 and 15 days newborn"
/lab_hosts="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_GHO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail is CGAAGTCAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.8%; Score 29; DB 14; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CTGAGTACCCAGAGAGAGCTGGCTGC 572
|||||
Db 535 CTGAGTACCCAGAGAGAGCTGGCTGC 563
|||||

RESULT 100

BQ442126
LOCUS
DEFINITION UI-M-EXO-bwy-n-04-0-UI.r1 NIH BMAP_EXO Mus musculus cDNA clone
IMAGE:5704899 5', mRNA sequence.
ACCESSION BQ442126
VERSION
KEYWORDS
SOURCE EST. GI:21245238
Mus musculus (house mouse)
ORGANISM
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

source
Seq primer: pYX-5.
Location/Qualifiers
1..692
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5704899"
/tissue_type="Whole brain"
/dev_stages="embryo 15.5 dpc"
/lab_hosts="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_EXO"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.8%; Score 29; DB 13; Length 692;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

544 CTGAGTACCCACAGGAGAGCTGGCTGC 572
572 CTGAGTACCCACAGGAGAGCTGGCTGC 600

Search completed: June 9, 2004, 00:49:37
Job time : 4002.48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 13:16:30 ; Search time 639.945 Seconds
(without alignments)
11676.867 Million cell updates/sec

Title: US-10-023-523-46
Perfect score: 1638
Sequence: 1 atgaagaacaaagacaaaaa.....agccacacctccgaggggcc 1638

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 2995936 seqs, 2280998010 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1638	100.0	1638	9	US-09-976-740-46
2	1638	100.0	1638	13	US-10-671-242-46
3	1638	100.0	1638	14	US-10-023-523-46
4	1638	100.0	1638	14	US-10-023-523-46
5	1638	100.0	1638	15	US-10-616-187-46
6	1536	93.8	2523	13	US-10-976-774-784
7	1439	87.9	4697	9	US-09-962-055-17
8	1439	87.9	4697	9	US-09-976-740-17
9	1439	87.9	4697	13	US-10-671-242-17
10	1439	87.9	4697	14	US-10-023-523-17
11	1439	87.9	4697	14	US-10-023-523-17
12	1439	87.9	4697	15	US-10-616-187-17
13	433	25.2	529	14	US-10-040-739-869
14	336	20.5	22255	9	US-09-976-740-51

15	336	20.5	22255	13	US-10-671-242-51	Sequence 51, Appl
16	336	20.5	22255	14	US-10-023-529-51	Sequence 51, Appl
17	336	20.5	22255	14	US-10-023-523-51	Sequence 51, Appl
18	336	20.5	22255	16	US-10-616-187-51	Sequence 51, Appl
19	225	2194	16	US-10-108-260A-1586	Sequence 1586, Ap	
20	140	8.5	405	9	US-09-764-869-467	Sequence 467, App
21	140	8.5	405	15	US-10-091-504-467	Sequence 467, App
22	140	8.5	405	16	US-10-227-577-467	Sequence 467, App
23	59	3.6	4722	9	US-09-962-055-14	Sequence 14, Appl
24	59	3.6	4722	9	US-09-976-740-14	Sequence 14, Appl
25	59	3.6	4722	13	US-10-671-242-14	Sequence 14, Appl
26	59	3.6	4722	14	US-10-023-529-14	Sequence 14, Appl
27	59	3.6	4722	14	US-10-023-523-14	Sequence 14, Appl
28	59	3.6	4722	16	US-10-616-187-14	Sequence 14, Appl
29	49	3.0	492	9	US-09-833-381-1334	Sequence 1334, Ap
30	37	2.3	937	12	US-09-876-143-1454	Sequence 1454, Ap
31	23	1.4	1351	13	US-10-425-114-6003	Sequence 6003, Ap
32	23	1.4	1955	13	US-10-425-114-32431	Sequence 32431, A
33	21	1.3	21	9	US-09-962-055-42	Sequence 42, Appl
34	21	1.3	21	9	US-09-976-740-42	Sequence 42, Appl
35	21	1.3	21	13	US-10-671-242-42	Sequence 42, Appl
36	21	1.3	21	14	US-10-023-529-42	Sequence 42, Appl
37	21	1.3	21	14	US-10-023-523-42	Sequence 42, Appl
38	21	1.3	21	16	US-10-616-187-42	Sequence 42, Appl
39	21	1.3	145	16	US-10-125-968-417	Sequence 417, App
40	21	1.3	160	13	US-10-085-783A-546	Sequence 546, App
41	21	1.3	160	16	US-10-242-535A-546	Sequence 546, App
42	21	1.3	261	13	US-10-085-783A-11558	Sequence 11558, A
43	21	1.3	261	16	US-10-242-535A-11558	Sequence 11558, A
44	21	1.3	340	13	US-10-085-783A-3476	Sequence 3476, Ap
45	21	1.3	340	16	US-10-242-535A-3476	Sequence 3476, Ap
46	21	1.3	393	9	US-09-920-300A-1581	Sequence 1581, Ap
47	21	1.3	393	14	US-10-033-528-1581	Sequence 1581, Ap
48	21	1.3	393	15	US-10-099-926-1581	Sequence 1581, Ap
49	21	1.3	400	13	US-10-085-783A-9227	Sequence 9227, Ap
50	21	1.3	400	16	US-10-242-535A-9227	Sequence 9227, Ap
51	21	1.3	403	15	US-10-198-846-6336	Sequence 6336, Ap
52	21	1.3	412	10	US-09-814-353-8962	Sequence 8962, Ap
53	21	1.3	416	15	US-10-198-846-9365	Sequence 9365, Ap
54	21	1.3	432	10	US-09-814-353-14275	Sequence 14275, A
55	21	1.3	446	9	US-09-777-564-423	Sequence 423, App
56	21	1.3	446	15	US-10-015-219-423	Sequence 423, App
57	21	1.3	447	9	US-09-777-564-139	Sequence 139, App
58	21	1.3	447	15	US-10-015-219-139	Sequence 139, App
59	21	1.3	458	9	US-09-920-455-171	Sequence 171, App
60	21	1.3	462	10	US-09-814-353-14284	Sequence 14284, A
61	21	1.3	497	10	US-09-814-353-15346	Sequence 15346, A
62	21	1.3	506	9	US-09-895-828-75	Sequence 75, Appl
63	21	1.3	506	15	US-10-114-666-75	Sequence 75, Appl
64	21	1.3	538	10	US-09-918-995-31626	Sequence 31626, A
65	21	1.3	538	9	US-09-813-358-167	Sequence 167, App
66	21	1.3	608	10	US-09-997-279-167	Sequence 167, App
67	21	1.3	631	15	US-10-002-631C-81	Sequence 81, Appl
68	21	1.3	824	9	US-09-822-849A-69	Sequence 69, Appl
69	21	1.3	857	9	US-09-822-849A-68	Sequence 68, Appl
70	21	1.3	893	9	US-09-880-107-2377	Sequence 2377, Ap
71	21	1.3	893	10	US-09-960-706-697	Sequence 697, App
72	21	1.3	893	10	US-09-873-319-439	Sequence 439, App
73	21	1.3	893	10	US-09-925-301-481	Sequence 481, App
74	21	1.3	1119	9	US-10-264-049-799	Sequence 799, App
75	21	1.3	1131	16	US-10-198-846-13193	Sequence 13193, A
76	21	1.3	1286	15	US-09-960-352-10281	Sequence 10281, A
77	20	1.2	328	9	US-10-027-632-290731	Sequence 290731, A
78	20	1.2	570	13	US-10-027-632-290731	Sequence 290731, A
79	20	1.2	576	16	US-10-027-632-290731	Sequence 1855, Ap
80	20	1.2	576	9	US-09-938-842A-1855	Sequence 1855, Ap
81	20	1.2	576	11	US-09-938-842A-1855	Sequence 1855, Ap
82	20	1.2	2045	9	US-09-960-253-172	Sequence 172, App
83	20	1.2	2045	16	US-10-115-831-34	Sequence 34, Appl
84	20	1.2	20029	9	US-09-764-847-1624	Sequence 1624, Ap
85	20	1.2	20029	15	US-10-092-154-1624	Sequence 1624, Ap
86	20	1.2	96599	16	US-10-085-117-100	Sequence 100, App
87	19	1.2	138	16	US-10-125-968-249	Sequence 249, App

c 88 19 1.2 217 10 US-09-814-353-1541 Sequence 1541, Ap
 c 89 19 1.2 217 10 US-09-814-353-7900 Sequence 7900, Ap
 90 19 1.2 230 10 US-09-814-353-1532 Sequence 1532, Ap
 91 19 1.2 230 10 US-09-814-353-7891 Sequence 7891, Ap
 92 19 1.2 472 13 US-10-085-783A-40305 Sequence 40305, A
 93 19 1.2 472 13 US-10-242-535A-40305 Sequence 40305, A
 94 19 1.2 519 15 US-10-198-846-9413 Sequence 9413, Ap
 95 19 1.2 522 13 US-10-027-632-181095 Sequence 181095,
 96 19 1.2 522 13 US-10-027-632-317786 Sequence 317786,
 97 19 1.2 522 16 US-10-027-632-181095 Sequence 181095,
 98 19 1.2 522 16 US-10-027-632-317786 Sequence 317786,
 99 19 1.2 541 13 US-10-027-632-285491 Sequence 285491,
 100 19 1.2 541 16 US-10-027-632-285491 Sequence 285491,

ALIGNMENTS

RESULT 1

US-09-976-740-46
 ; Sequence 46, Application US/09976740
 ; Publication No. US20020194633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/09/976,740
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: Past-Seq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 1638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1638)
 US-09-976-740-46

Query Match 100.0%; Score 1638; DB 9; Length 1638;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGAACCAAGACAAAGAGCGGGCTGCGCAACATCCAAATCCAAAGACGCCCA 60
 Db 1 ATGAGAACCAAGACAAAGAGCGGGCTGCGCAACATCCAAATCCAAAGACGCCCA 60
 QY 61 GGACAAACCGGAAGCAGGACCCGAGGAGCCGAGGAGCCGAGCGCGGCTCTGCA 120
 Db 61 GGACAAACCGGAAGCAGGACCCGAGGAGCCGAGGAGCCGAGCGCGGCTCTGCA 120
 QY 121 GTAGAGCAAGAGGTCCCGCAGCAGCAGGCTCTCGAGAGCCGAGGAGGTCTCAAGCC 180
 Db 121 GTAGAGCAAGAGGTCCCGCAGCAGCAGGCTCTCGAGAGCCGAGGAGGTCTCAAGCC 180
 QY 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCCCAACTGGAA 240
 Db 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCCCAACTGGAA 240
 QY 241 GACATACAGCACATCTGTGTGGACAAATAACAGGGGGGGGGCGGCGAGGATGGGCA 300

Db 241 GACATACAGCACATCTGTGTGGACAAATAACAGGGGGGGGGCGGCGAGGATGGGCA 300
 QY 301 CAGGGGTGAGCCGGCTGAACCCGAGAGTGCAGAGAAAGTCCCGGACCTTGTGGCAAGGAAT 360
 Db 301 CAGGGGTGAGCCGGCTGAACCCGAGAGTGCAGAGAAAGTCCCGGACCTTGTGGCAAGGAAT 360
 QY 361 GGGGAGCCCTGAACCAACTCCAGTACTCAATGGAGAGAAAGAACCTTCAAGGGGATCA 420
 Db 361 GGGGAGCCCTGAACCAACTCCAGTACTCAATGGAGAGAAAGAACCTTCAAGGGGATCA 420
 QY 421 AACACAGAGAGATCCGCGAGAGTCCAGAGGTCCGAGACCCGAGACCTATCGAAGGCCACAG 480
 Db 421 AACACAGAGAGATCCGCGAGAGTCCAGAGGTCCGAGACCCGAGACCTATCGAAGGCCACAG 480
 QY 481 GAGAGAGAGAGATCCGCGAGAGTCCAGAGGTCCGAGACCTATCGAAGGCCACAG 540
 Db 481 GAGAGAGAGAGATCCGCGAGAGTCCAGAGGTCCGAGACCTATCGAAGGCCACAG 540
 QY 541 ACTCTGAGTACCCAGAGAGAGTGGCTGCTCTGTGCAAGAGTATGCTGAACTGCTG 600
 Db 541 ACTCTGAGTACCCAGAGAGAGTGGCTGCTCTGTGCAAGAGTATGCTGAACTGCTG 600
 QY 601 GAGGAGCACCGGAATTCAAGAGCAGATGAAGCTCTTACAGAGAAAGACGAGCCAGCTG 660
 Db 601 GAGGAGCACCGGAATTCAAGAGCAGATGAAGCTCTTACAGAGAAAGACGAGCCAGCTG 660
 QY 661 GTGCAAGAGAGGACCACTCGCGGTGAGCAGACAGAGGCGCTCTGGCCGAGCAAG 720
 Db 661 GTGCAAGAGAGGACCACTCGCGGTGAGCAGACAGAGGCGCTCTGGCCGAGCAAG 720
 QY 721 CTTGAGAGCCTATGCGGTGAGCTGCGCGGACCAACCGCTCCCTCAAGAGAAAGGTGTG 780
 Db 721 CTTGAGAGCCTATGCGGTGAGCTGCGCGGACCAACCGCTCCCTCAAGAGAAAGGTGTG 780
 QY 781 CAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 Db 781 CAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 QY 841 CTGAATGACATTCAGCTGACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 900
 Db 841 CTGAATGACATTCAGCTGACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 900
 QY 901 GAGAAATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTTGAGCAGATGAGCTGCGGAG 960
 Db 901 GAGAAATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTTGAGCAGATGAGCTGCGGAG 960
 QY 961 GAGCATATCGACAAAGTCTTCAAAACAGAGGACCTTCAAAACAGAGGAGGAGGAGGAGGAG 1020
 Db 961 GAGCATATCGACAAAGTCTTCAAAACAGAGGACCTTCAAAACAGAGGAGGAGGAGGAGGAG 1020
 QY 1021 CTCAGAGGCGCCAGGAGATCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 Db 1021 CTCAGAGGCGCCAGGAGATCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 QY 1081 TTTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
 Db 1081 TTTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
 QY 1141 ACCACCTTGAAGCAACAGCTTCCCTTATACACAGAGAGTTTGAAGGAGTTTCCAGAACACA 1200
 Db 1141 ACCACCTTGAAGCAACAGCTTCCCTTATACACAGAGAGTTTGAAGGAGTTTCCAGAACACA 1200
 QY 1201 CTTTCCAAAGCAGGAGGATTTCCACCATTCAGCAGGAGATGGAAGAGTACTAAG 1260
 Db 1201 CTTTCCAAAGCAGGAGGATTTCCACCATTCAGCAGGAGATGGAAGAGTACTAAG 1260
 QY 1261 AAGATCAAGAGCTGGAGAGAAACCAATGATACCGGTCCCGGTGGGAGAGCAGCAAC 1320
 Db 1261 AAGATCAAGAGCTGGAGAGAAACCAATGATACCGGTCCCGGTGGGAGAGCAGCAAC 1320
 QY 1321 AAGCCCTGCTTGAAGTGGTGGAGAGAAACAGTCCGGGATTAAGAACTGGAGGGCTG 1380

Db 1321 AAGGCCCTGCTTGAGATGGCTGAGGAGAAAAACAGTCCGGGATAAAAGAACTGGAGGCGCTG 1380
QY 1381 CAGGTAAATCAACGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1440
Db 1381 CAGGTAAATCAACGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1440
QY 1441 CTGAACAAGAGGGTACAGGACCTGAGTGTGTGTGGCCAGGGTCCCTCACTACAGTGGC 1500
Db 1441 CTGAACAAGAGGGTACAGGACCTGAGTGTGTGTGGCCAGGGTCCCTCACTACAGTGGC 1500
QY 1501 CTTGAGAGAGGCGCAGAGGGCTTGGGCTCAAGCACCCAGCTCCCGCAGGGTCAAGAA 1560
Db 1501 CTTGAGAGAGGCGCAGAGGGCTTGGGCTCAAGCACCCAGCTCCCGCAGGGTCAAGAA 1560
QY 1561 GCGCCTTGTCTACCGAGAGCAGCGAGCACAGAGAGCATCAGGCTGGGCTCAAGAG 1620
Db 1561 GCGCCTTGTCTACCGAGAGCAGCGAGCACAGAGAGCATCAGGCTGGGCTCAAGAG 1620
QY 1621 CCACCTCGCCAGGGCC 1638
Db 1621 CCACCTCGCCAGGGCC 1638

RESULT 2

US-10-671-242-46
; Sequence 46, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1638)
US-10-671-242-46

Query Match 100.0%; Score 1638; DB 13; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGAACCCAGAACAGCGGGCTGCCAAACAAATCCCAATCCAAAAGCAGCCCA 60
Db 1 ATGAAGAACCCAGAACAGCGGGCTGCCAAACAAATCCCAATCCAAAAGCAGCCCA 60
QY 61 GGACAAACCGAGCAGGACCCGAGGAGCCCGAGAGCGGGCCAGCCAGGGGGTCTCTGCA 120
Db 61 GGACAAACCGAGCAGGACCCGAGGAGCCCGAGAGCGGGCCAGCCAGGGGGTCTCTGCA 120
QY 121 GTAGAGCAGAGTCCCGGAGCAGCGAGGCTCTCGAAGCCGGAGGGTGTCTCAAGCC 180

Db 121 GTAGAGCAGAGTCCCGGAGCAGCGAGGCTCTCGAAGCCGGAGGGTGTCTCAAGCC 180
QY 181 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCCGCAACTGGAA 240
Db 181 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCCGCAACTGGAA 240
QY 241 CACATCTAGCACAATCTGTGTGGAACAATACAGAGGGGGCCCGGGAGGATGGGGCA 300
Db 241 CACATCTAGCACAATCTGTGTGGAACAATACAGAGGGGGCCCGGGAGGATGGGGCA 300
QY 301 CAGGTGAGCGGCTGAACCCGAGAGTGCAGAGAGTCCCGGAGCTATGTGGCAAGGAAT 360
Db 301 CAGGTGAGCGGCTGAACCCGAGAGTGCAGAGAGTCCCGGAGCTATGTGGCAAGGAAT 360
QY 361 GGGGAGCCTGAACCAACTCCAGTGTCAATCGAGAGAGGAACCCCTCAAGGGGGATCCA 420
Db 361 GGGGAGCCTGAACCAACTCCAGTGTCAATCGAGAGAGGAACCCCTCAAGGGGGATCCA 420
QY 421 AACACAGAGAGATCCCGCAGAGTGCAGAGTCCGAGACCCGAGACCATCGAAGGCCACAG 480
Db 421 AACACAGAGAGATCCCGCAGAGTGCAGAGTCCGAGACCCGAGACCATCGAAGGCCACAG 480
QY 481 GAGAAAGAAAAAGCCAGGGTTTGGGTAAAGGAGATCAGTTTGTCTGATGCAGACATTGAAT 540
Db 481 GAGAAAGAAAAAGCCAGGGTTTGGGTAAAGGAGATCAGTTTGTCTGATGCAGACATTGAAT 540
QY 541 ACTCTGAGTACCCAGAGAGAGAGTGTGCTGCTGTGTCAGAGAGTATGCTGAAGTGTG 600
Db 541 ACTCTGAGTACCCAGAGAGAGAGTGTGCTGCTGTGTCAGAGAGTATGCTGAAGTGTG 600
QY 601 GAGGAGACCCGAATTACAGAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
Db 601 GAGGAGACCCGAATTACAGAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
QY 661 GTGCAAGAGAGGACCACTGCGCGGTGAGCAGACAGAGGCGCTCTGCGCCGCGAGCAAG 720
Db 661 GTGCAAGAGAGGACCACTGCGCGGTGAGCAGACAGAGGCGCTCTGCGCCGCGAGCAAG 720
QY 721 CTTGAGAGCTATGCGGTGAGCTGAGCGGACAGCCGCTCCCTCAAGAGAGAGTGTG 780
Db 721 CTTGAGAGCTATGCGGTGAGCTGAGCGGACAGCCGCTCCCTCAAGAGAGAGTGTG 780
QY 781 CAGCGGGCCCGGGAGGAGGAGAGGCAAGGAGGTGACTCTGCACTTCAGGTGACA 840
Db 781 CAGCGGGCCCGGGAGGAGGAGAGGCAAGGAGGTGACTCTGCACTTCAGGTGACA 840
QY 841 CTGAATGACATTCAGCTGAGTGAACAGACACATGAGCGCACTCCAAAGTGGCCAA 900
Db 841 CTGAATGACATTCAGCTGAGTGAACAGACACATGAGCGCACTCCAAAGTGGCCAA 900
QY 901 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGCGAG 960
Db 901 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGCGAG 960
QY 961 GAGCATATCGACAAAGTCTTCAAAACAGAGGACCTTACAAACAGCAGCTGGTGGATGCCAAG 1020
Db 961 GAGCATATCGACAAAGTCTTCAAAACAGAGGACCTTACAAACAGCAGCTGGTGGATGCCAAG 1020
QY 1021 CTTCCAGCGGGCCAGGAGTGTCTAAAGGGCAGAGAGCGGCACAGCGGAGAGGAT 1080
Db 1021 CTTCCAGCGGGCCAGGAGTGTCTAAAGGGCAGAGAGCGGCACAGCGGAGAGGAT 1080
QY 1081 TTTTCTCTGAAAGAGGCGAGTAGAGTCCCAGAGGATGTGTGAGCTGATGAAGCAGCAAG 1140
Db 1081 TTTTCTCTGAAAGAGGCGAGTAGAGTCCCAGAGGATGTGTGAGCTGATGAAGCAGCAAG 1140
QY 1141 ACCACCTGAAGCAACAGCTTCCCTTATACAGAGAGAGTTTGAGGATTCAGAGACACA 1200
Db 1141 ACCACCTGAAGCAACAGCTTCCCTTATACAGAGAGAGTTTGAGGATTCAGAGACACA 1200
QY 1201 CTTTCCAAAGCAGCGAGGTATTCCACCATTCAGAGAGGATGGAAGAGTACTTAAG 1260
Db 1201 CTTTCCAAAGCAGCGAGGTATTCCACCATTCAGAGAGGATGGAAGAGTACTTAAG 1260

QY 1261 AAGATCAAGAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
DB 1261 AAGATCAAGAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
QY 1321 AAGCCCTGCTTGAGATGGCTGAGAGAGAAACAGTCCGGGATAAAGAACTGGAGGGCTG 1380
DB 1321 AAGCCCTGCTTGAGATGGCTGAGAGAGAAACAGTCCGGGATAAAGAACTGGAGGGCTG 1380
QY 1381 CAGGTAAATCCACCGCTGGAGAGCTGTGCGGCACATGACAGACAGAGCCCATGAC 1440
DB 1381 CAGGTAAATCCACCGCTGGAGAGCTGTGCGGCACATGACAGACAGAGCCCATGAC 1440
QY 1441 CTGAACAGAGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
DB 1441 CTGAACAGAGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
QY 1501 CCTGAGAGAGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCGCCAGGGTCAAGAA 1560
DB 1501 CCTGAGAGAGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCGCCAGGGTCAAGAA 1560
QY 1561 GCGCTTGTCTACCCAGAGAGCAGGACAGACAGAGCATCAGCCAGAGCTGGGGCTCAAGAG 1620
DB 1561 GCGCTTGTCTACCCAGAGAGCAGGACAGACAGAGCATCAGCCAGAGCTGGGGCTCAAGAG 1620
QY 1621 CCCACCTCCGCGAGGGCC 1638
DB 1621 CCCACCTCCGCGAGGGCC 1638

RESULT 3

US-10-023-529-46
; Sequence 46, Application US/10023529
; Publication No. US2002012388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-10-023-529-46

Query Match 100.0%; Score 1638; DB 14; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGAACCAAGACAAAGAGAGCGGGCTGCCAACAATCCCAATCCAAAGACGCCCA 60
DB 1 ATGAAGAACCAAGACAAAGAGAGCGGGCTGCCAACAATCCCAATCCAAAGACGCCCA 60

QY 61 GCACAACCGAAGCAGGACCCGAGGGAGCCGAGGAGCGGCCAGCCAGCGCGCTCTCTGCA 120
DB 61 GCACAACCGAAGCAGGACCCGAGGGAGCCGAGGAGCGGCCAGCCAGCGCGCTCTCTGCA 120
QY 121 GTAGAAGCAGAAAGTCCCGGAGCAGCAGGCTCTCTCGAAGCCGAGAGGGTCTCAAGCC 180
DB 121 GTAGAAGCAGAAAGTCCCGGAGCAGCAGGCTCTCTCGAAGCCGAGAGGGTCTCAAGCC 180
QY 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTATGTCCTGAGGAGCTGAGCGCCAACTGGAA 240
DB 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTATGTCCTGAGGAGCTGAGCGCCAACTGGAA 240
QY 241 GACATACCTGAGCACAATCTGTGTGGACAATAACCAAGGGGGCCCCCGGAGAGTGGGGCA 300
DB 241 GACATACCTGAGCACAATCTGTGTGGACAATAACCAAGGGGGCCCCCGGAGAGTGGGGCA 300
QY 301 CAGGCTGAGCCGGCTGAACCCGAGATGACAGAAAGTCCCGGACCTATGTGGCAAGAAAT 360
DB 301 CAGGCTGAGCCGGCTGAACCCGAGATGACAGAAAGTCCCGGACCTATGTGGCAAGAAAT 360
QY 361 GGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAGGAAACCTCCAGGGGGATCCA 420
DB 361 GGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAGGAAACCTCCAGGGGGATCCA 420
QY 421 AACACAGAGAGATCCGGCAGAGTGAAGGCTCGGAGACCGAGACCAATCGAAGGCCACAG 480
DB 421 AACACAGAGAGATCCGGCAGAGTGAAGGCTCGGAGACCGAGACCAATCGAAGGCCACAG 480
QY 481 GAGAGAGAGAGCCAGGGTTTGGGTAAGGAGATCAGTTGCTGATGACACATTTGAAT 540
DB 481 GAGAGAGAGAGCCAGGGTTTGGGTAAGGAGATCAGTTGCTGATGACACATTTGAAT 540
QY 541 ACTCTGAGTACCCAGAGAGAGAGTGGCTCTCTGTGCAAGAGATGCTGAACTGCTG 600
DB 541 ACTCTGAGTACCCAGAGAGAGAGTGGCTCTCTGTGCAAGAGATGCTGAACTGCTG 600
QY 601 GAGGAGCACCAGATTCACAGAGCAGATGAAGCTCTTACAGAAAAAGCAGAGCCAGCTG 660
DB 601 GAGGAGCACCAGATTCACAGAGCAGATGAAGCTCTTACAGAAAAAGCAGAGCCAGCTG 660
QY 661 GTCAAGAGAGAGGACCACTCGCGGTGAGCAGCAGCAAGGCGCTCCTGGCCCGCAGCAAG 720
DB 661 GTCAAGAGAGAGGACCACTCGCGGTGAGCAGCAGCAGCAAGGCGCTCCTGGCCCGCAGCAAG 720
QY 721 CTTGAGACCTATGCGGTGAGCTGAGCGGCACAAACCGCTCCCTCAAGAGAAAGGTGTG 780
DB 721 CTTGAGACCTATGCGGTGAGCTGAGCGGCACAAACCGCTCCCTCAAGAGAAAGGTGTG 780
QY 781 CAGCGGGCCCGGAGGAGGAGAGCGCAAGGAGTGAACCTCGCACTTCCAGGTGACA 840
DB 781 CAGCGGGCCCGGAGGAGGAGAGCGCAAGGAGTGAACCTCGCACTTCCAGGTGACA 840
QY 841 CTGAATGATTCAGTTCAGATGCAAGCAGCAATGAGCGCAACTCCAGCTCGCCAA 900
DB 841 CTGAATGATTCAGTTCAGATGCAAGCAGCAATGAGCGCAACTCCAGCTCGCCAA 900
QY 901 GAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGTATGAGCTGCGGAG 960
DB 901 GAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGTATGAGCTGCGGAG 960
QY 961 GAGCATATCGACAAAGTCTTCAAAACACAGGACCTCAACAGCAGAGCTGGTGGATGCCAAG 1020
DB 961 GAGCATATCGACAAAGTCTTCAAAACACAGGACCTCAACAGCAGAGCTGGTGGATGCCAAG 1020
QY 1021 CTCAGCAGGCCCCAGGAGATGCTTAAGGAGCAGAGCGGCCACAGCGGAGAGGAT 1080
DB 1021 CTCAGCAGGCCCCAGGAGATGCTTAAGGAGCAGAGCGGCCACAGCGGAGAGGAT 1080
QY 1081 TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGCTGAGCTGATGAAGCAGCAAGAG 1140
DB 1081 TTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGCTGAGCTGATGAAGCAGCAAGAG 1140

901 GAGACATGGAGTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGGCGAG 960
961 GAGCATATGACAAAGTCTTCAACACAAAGGACCTACAAACAGCAGCTGGTGGTGGTCAAG 1020
961 GAGCATATGACAAAGTCTTCAACACAAAGGACCTACAAACAGCAGCTGGTGGTGGTCAAG 1020
1021 CTCACAGGCCCCAGGAGATGCTTAAAGAGGCGAAGAGCGGACACAGCGGAGAGGAT 1080
1021 CTCACAGGCCCCAGGAGATGCTTAAAGAGGCGAAGAGCGGACACAGCGGAGAGGAT 1080
1081 TTTTCTCTGAAAGAGGACGCTAGAGTCCACAGAGGATGTTGAGCTGATGAGCAGCAAGAG 1140
1081 TTTTCTCTGAAAGAGGACGCTAGAGTCCACAGAGGATGTTGAGCTGATGAGCAGCAAGAG 1140
1141 ACCCACTGAAAGCAACAGCTTGGCTTATACACAGAGAGTTTGGAGGTTCCAGAACACA 1200
1141 ACCCACTGAAAGCAACAGCTTGGCTTATACACAGAGAGTTTGGAGGTTCCAGAACACA 1200
1201 CTTTCCAAAAGAGCAGGATTTTCAACCACTTCAACAGAGAGTGAAGAGATGACTAAG 1260
1201 CTTTCCAAAAGAGCAGGATTTTCAACCACTTCAACAGAGAGTGAAGAGATGACTAAG 1260
1261 AAGATCAAGAGGCTGGAGAAAGAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
1261 AAGATCAAGAGGCTGGAGAAAGAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
1321 AAGGCTCTGTTGAGATGGTGAAGAGGAGAAACAGTCCGGGATTAAGAACTGGAGGCGCTG 1380
1321 AAGGCTCTGTTGAGATGGTGAAGAGGAGAAACAGTCCGGGATTAAGAACTGGAGGCGCTG 1380
1381 CAGGTAAATCCAAAGGCTGGAGAGGCTGTGCGGGGCACTGACAGACAGAGCGCAATGAC 1440
1381 CAGGTAAATCCAAAGGCTGGAGAGGCTGTGCGGGGCACTGACAGACAGAGCGCAATGAC 1440
1441 CTGAAACAGAGGTTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
1441 CTGAAACAGAGGTTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
1501 CTTGAGAGGAGGCGAGAGGCGCTGGGGCTCAAGCACCAGCTCCCGCAGGGTCAAGAA 1560
1501 CTTGAGAGGAGGCGAGAGGCGCTGGGGCTCAAGCACCAGCTCCCGCAGGGTCAAGAA 1560
1561 GCGCTTGTCTACCCAGGAGCACCGAGCACAGAGAGCATCAGGAGGATGAGGCTCAAGAG 1620
1561 GCGCTTGTCTACCCAGGAGCACCGAGCACAGAGAGCATCAGGAGGATGAGGCTCAAGAG 1620
1621 CCCACTCCGCGAGGCGC 1638
1621 CCCACTCCGCGAGGCGC 1638

RESULT 6
US-10-276-774-784
; Sequence 784, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10276, 774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 784
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-784

Query Match 93.8%; Score 1536; DB 13; Length 2523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGACCAAGCAAAAGACGGGGCTGCCAAACAAATCCAAATCCAAAGAGCAGCCCA 60
DB 119 ATGAGAACCAAGCAAAAGACGGGGCTGCCAAACAAATCCAAATCCAAAGAGCAGCCCA 178
QY 61 GGCACACCGGAGAGGACCCGAGGAGCCAGGAGGCGGCCAGCCAGGCGGCTCTTGCA 120
DB 179 GGCACACCGGAGAGGACCCGAGGAGCCAGGAGGCGGCCAGCCAGGCGGCTCTTGCA 238
QY 121 GTAGAGCAGAGAGTCCCGGAGCAGCAGCAGCTCTCGGAAGCGGAGGCTGCTCAAGCC 180
DB 239 GTAGAGCAGAGAGTCCCGGAGCAGCAGCAGCTCTCGGAAGCGGAGGCTGCTCAAGCC 298
QY 181 AGAACGGCTCAGTCTGGGGCCCTTCTGATGATCTCTGAGAGCTGAGCGGCCCAACTGGA 240
DB 299 AGAACGGCTCAGTCTGGGGCCCTTCTGATGATCTCTGAGAGCTGAGCGGCCCAACTGGA 358
QY 241 GACATACCTGAGCAGATCTGTGGAACATTAACAGAGGCGGCCCGGCGAGGATGGGCA 300
DB 359 GACATACCTGAGCAGATCTGTGGAACATTAACAGAGGCGGCCCGGCGAGGATGGGCA 418
QY 301 CAGGTGAGCGGCTGACCCGAGAGTGCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 360
DB 419 CAGGTGAGCGGCTGACCCGAGAGTGCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 478
QY 361 GGGGAGCCTGAAACCACTCCAGTAGTCAATGAGAGAGAGAACCTTCAAGGGGATCCA 420
DB 479 GGGGAGCCTGAAACCACTCCAGTAGTCAATGAGAGAGAGAACCTTCAAGGGGATCCA 538
QY 421 AACACAGAGAGATCCCGCAGAGTGCAGGTCGGAGAGCCGAGACCTCGAGGCGCACAG 480
DB 539 AACACAGAGAGATCCCGCAGAGTGCAGGTCGGAGAGCCGAGACCTCGAGGCGCACAG 598
QY 481 GAGAGAAAAAGCCAGCGTTTGGGTAAAGAGATCACTTGTGATGATGAGACATTTGAT 540
DB 599 GAGAGAAAAAGCCAGGTTTGGGGAAGAGATCACGTTGCTGATGAGACATTTGAT 658
QY 541 ACTCTGAGTACCCAGAGAGAGTGGCTGCTGTGTGCAAGAGTATGCTGAATGCTG 600
DB 659 ACTCTGAGTACCCAGAGAGAGTGGCTGCTGTGTGCAAGAGTATGCTGAATGCTG 718
QY 601 GAGGAGACCGGAATTCACAGAGCAGATGAAGCTCTACAGAAAAAGCAGAGCAGCTG 660
DB 719 GAGGAGACCGGAATTCACAGAGCAGATGAAGCTCTACAGAAAAAGCAGAGCAGCTG 778
QY 661 GTGCAAGAGAGGACCCACTGCGCGGTGAGCAGCAGCAAGGCGCTCTGCGCGCGCAGCAAG 720
DB 779 GTGCAAGAGAGGACCCACTGCGCGGTGAGCAGCAGCAAGGCGCTCTGCGCGCGCAGCAAG 838
QY 721 CTTGAGAGCCTATGCGGTGAGCTGAGCGGACACACCGCTCCCTCAAGAGAGAGGCTG 780
DB 839 CTTGAGAGCCTATGCGGTGAGCTGAGCGGACACACCGCTCCCTCAAGAGAGAGGCTG 898
QY 781 CAGCGCGGCCCGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 899 CAGCGCGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
QY 841 CTGAATGACATTCAGCTGAGATGGAACAGACATGAGCGCACTCAAGTGGCGGCA 900
DB 959 CTGAATGACATTCAGCTGAGATGGAACAGACATGAGCGCACTCAAGTGGCGGCA 1018
QY 901 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGGCGGAG 960
DB 1019 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGGCGGAG 1078
QY 961 GAGCATATCGACAAAGTCTTCAAAACAGAGGACCTTACAAACAGAGGAGCTGGTGGATGCCAAG 1020
DB 1079 GAGCATATCGACAAAGTCTTCAAAACAGAGGAGCTTACAAACAGAGGAGCTGGTGGATGCCAAG 1138


```
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/10/023,523
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-523-51

Query Match 20.5%; Score 336; DB 14; Length 22255;
Best Local Similarity 100.0%; Pred. No. 3.9e-159;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 231
Db 2999 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 3058
QY 232 CAACCTGGAAGACATACCTGAGCACATCTGTGTGACAAATACCAAGGGGGCCCGGCGAG 291
Db 3059 CAACCTGGAAGACATACCTGAGCACATCTGTGTGACAAATACCAAGGGGGCCCGGCGAG 3118
QY 292 GATGGGACACAGGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 351
Db 3119 GATGGGACACAGGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 411
Db 3179 GCAAGGAATGGGAGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 3298
QY 472 AGGCCACAGGAGAGAAAGAAAGCCAGGGTTGGGT 507
Db 3299 AGGCCACAGGAGAGAAAGAAAGCCAGGGTTGGGT 3334

RESULT 17
US-10-023-523-51
; Sequence 51, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 1997-06-03
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 51
; LENGTH: 22255
```

```
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-523-51

Query Match 20.5%; Score 336; DB 14; Length 22255;
Best Local Similarity 100.0%; Pred. No. 3.9e-159;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 231
Db 2999 GCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATCTCTCTGAGGAGCTGAGCCGC 3058
QY 232 CAACCTGGAAGACATACCTGAGCACATCTGTGTGACAAATACCAAGGGGGCCCGGCGAG 291
Db 3059 CAACCTGGAAGACATACCTGAGCACATCTGTGTGACAAATACCAAGGGGGCCCGGCGAG 3118
QY 292 GATGGGACACAGGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 351
Db 3119 GATGGGACACAGGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 411
Db 3179 GCAAGGAATGGGAGCTGAGCCGGCTGAACCCGAGATGACAGAGTCCCGGACCTATGTG 3298
QY 472 AGGCCACAGGAGAGAAAGAAAGCCAGGGTTGGGT 507
Db 3299 AGGCCACAGGAGAGAAAGAAAGCCAGGGTTGGGT 3334

RESULT 18
US-10-616-187-51
; Sequence 51, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 51
; LENGTH: 22255
```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-616-187-51

Query Match 20.5%; Score 336; DB 16; Length 22255;
Best Local Similarity 100.0%; Pred. No. 3.9e-159;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCTCAAGCCAGACGAGCTCAGTCTGGGGCCCTTCGTGATCTCTCAGGAGCTGAGCCGC 231
DB 2999 GCTCAAGCCAGACGAGCTCAGTCTGGGGCCCTTCGTGATCTCTCAGGAGCTGAGCCGC 3058
QY 232 CAACCTGAAGACATATCTGAGCACAATCTGTGTGGACAATAACACAGGGGGCCCCGGCAG 291
DB 3059 CAACCTGAAGACATATCTGAGCACAATCTGTGTGGACAATAACACAGGGGGCCCCGGCAG 3118
QY 292 GATGGGGCACAGGTGAGCCGGCTGAACCCGAGATGCGAGAGTCCCGGACCTATGTG 351
DB 3119 GATGGGGCACAGGTGAGCCGGCTGAACCCGAGATGCGAGAGTCCCGGACCTATGTG 3178
QY 352 GCAAGGAATGGGAGCCTGACCAACTCCAGTAGTCAATGGAGAGGAAGAACCTCCAAAG 411
DB 3179 GCAAGGAATGGGAGCCTGACCAACTCCAGTAGTCAATGGAGAGGAAGAACCTCCAAAG 3238
QY 412 GGGGATCCAAACACAGAGAGATCCGCGACAGTACGAGTCCGAGACCGAGACCATCGA 471
DB 3239 GGGGATCCAAACACAGAGAGATCCGCGACAGTACGAGTCCGAGACCGAGACCATCGA 3298
QY 472 AGGCCACAGGAGAGAAAAAGCAAGGGTTTGGGT 507
DB 3299 AGGCCACAGGAGAGAAAAAGCAAGGGTTTGGGT 3334

RESULT 19
US-10-108-260A-1586
; Sequence 1586, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1586
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1586

Query Match 13.7%; Score 225; DB 16; Length 2194;
Best Local Similarity 99.4%; Pred. No. 5.5e-103;
Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGAACCAAGACAAAAAGACGGGGCTGCCAAATCCAAATCCAAAGACAGCCCA 60
DB 66 ATGAGAACCAAGACAAAAAGACGGGGCTGCCAAATCCAAATCCAAAGACAGCCCA 125
QY 61 GGAACCCGAGACAGACCCGAGGAGAGCCAGAGCGGCCAGCCAGCGGGCTCTCTGCA 120
DB 126 GGAACCCGAGACAGACCCGAGGAGAGCCAGAGCGGCCAGCGGGCTCTCTGCA 185
QY 121 GTGAAGCAGAAAGTCCCGGACGAGCAGCGCTCTCGGAAGCGGAGGGTGTCAAGCC 180
DB 186 GTGAAGCAGAAAGTCCCGGACGAGCAGCGCTCTCGGAAGCGGAGGGGTCTAAGCC 245
QY 181 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGTGAAGCCGACCTGGA 240
DB 246 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGTGAAGCCGACCTGGA 305
QY 241 GACATCTGACACATCTGTGTGGACAATAACACAGGGGGCCCCCGGAGGATGGGCA 300

DB 306 GACATACTGAGCACAATCTGTGTGGACAATAACACAGGGGGCCCCCGGAGGATGGGGCA 365
QY 301 CAGGGTGAGCGGGCTGAACCCGGAAGAT 327
DB 366 CAGGGTGAGCGGGCTGAACCCGGAAGAT 392

RESULT 20
US-09-764-869-467
; Sequence 467, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-467

Query Match 8.5%; Score 140; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.4e-60;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 GCCAACTGGAAGACATCTGAGCACAATCTGTGTGGACAATAACACAGGGGGCCCCCGCG 289
DB 88 GCCAACTGGAAGACATCTGAGCACAATCTGTGTGGACAATAACACAGGGGGCCCCCGCG 147
QY 290 AGGATGGGGCACAGGTCAGCGGCTGAACCCGAGATGCGAGAGTCCCGGACCTATG 349
DB 148 AGGATGGGGCACAGGTCAGCGGCTGAACCCGAGATGCGAGAGTCCCGGACCTATG 207
QY 350 TGGCAGGAATGGGAGCCT 369
DB 208 TGGCAGGAATGGGAGCCT 227

RESULT 21
US-10-091-504-467
; Sequence 467, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)

OTHER INFORMATION: n equals a,t,g, or c
US-10-031-504-467

Query Match 8.5%; Score 140; DB 15; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.4e-60;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 GCCAACTGGAAGACATAGTGGACATACCTGTGGCAATAACCCGAGGAGTCCCGGCG 289
DB 88 GCCAACTGGAAGACATAGTGGACATACCTGTGGCAATAACCCGAGGAGTCCCGGCG 147

QY 290 AGGATGGGGCCACAGGCTGAGCCGCTGAACCCGAGGAGTCCCGGACCTATG 349
DB 148 AGGATGGGGCCACAGGCTGAGCCGCTGAACCCGAGGAGTCCCGGACCTATG 207

QY 350 TGGCAAGGAATGGGAGCCT 369
DB 208 TGGCAAGGAATGGGAGCCT 227

RESULT 22

US-10-227-577-467
; Sequence 467, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26

; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 467

; LENGTH: 405
; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (377)

; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (398)

; OTHER INFORMATION: n equals a,t,g, or c

US-10-577-467

Query Match 8.5%; Score 140; DB 16; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.4e-60;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 GCCAACTGGAAGACATAGTGGACATACCTGTGGCAATAACCCGAGGAGTCCCGGCG 289
DB 88 GCCAACTGGAAGACATAGTGGACATACCTGTGGCAATAACCCGAGGAGTCCCGGCG 147

DB 88 GCCAACTGGAAGACATAGTGGACATACCTGTGGCAATAACCCGAGGAGTCCCGGCG 147
QY 290 AGGATGGGGCCACAGGCTGAGCCGCTGAACCCGAGGAGTCCCGGACCTATG 349
DB 148 AGGATGGGGCCACAGGCTGAGCCGCTGAACCCGAGGAGTCCCGGACCTATG 207
QY 350 TGGCAAGGAATGGGAGCCT 369
DB 208 TGGCAAGGAATGGGAGCCT 227

RESULT 23

US-09-962-055-14
; Sequence 14, Application US/09962055
; Patent No. US2002005203A1
; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4722 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 61...1731

; SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-962-055-14

Query Match 3.6%; Score 59; DB 9; Length 4722;

Best Local Similarity 100.0%; Pred. No. 3.6e-19;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGGAGAGCTGTGCGGCACTGCAGACAGCGCAATGACCTGAACAAGGGT 1454
DB 1456 CGGCTGGAGAGCTGTGCGGCACTGCAGACAGCGCAATGACCTGAACAAGGGT 1514

RESULT 24


```
US-09-976-740-14
; Sequence 14, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-09-976-740-14

Query Match      3.6%; Score 59; DB 9; Length 4722;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1454
DB 1456 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 25
US-10-671-242-14
; Sequence 14, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus

Query Match      3.6%; Score 59; DB 9; Length 4722;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1454
DB 1456 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 26
US-10-023-529-14
; Sequence 14, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-10-023-529-14

Query Match      3.6%; Score 59; DB 14; Length 4722;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1396 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1454
DB 1456 CGGCTGGAGAGCTGTGCCGGCACTGCACACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 27
US-10-023-523-14
; Sequence 14, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
```

;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)...(1731)
US-10-023-523-14

Query Match 3.6%; Score 59; DB 14; Length 4722;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1396 CGGCTGGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1454
Db 1456 CGGCTGGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 28
US-10-616-187-14
;; Sequence 14, Application US/10616187
;; Publication No. US20040013668A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; TITLE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/616,187
;; CURRENT FILING DATE: 2003-07-09
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)...(1731)
US-10-616-187-14

Query Match 3.6%; Score 59; DB 16; Length 4722;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1396 CGGCTGGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1454

Db 1456 CGGCTGGAGAGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1514
RESULT 29
US-09-833-381-1334
;; Sequence 1334, Application US/09833381
;; Patent No. US20020132090A1
;; GENERAL INFORMATION:
;; APPLICANT: Robison, Keith E.
;; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;; FILE REFERENCE: 5800-119
;; CURRENT APPLICATION NUMBER: US/09/833,381
;; CURRENT FILING DATE: 2001-04-11
;; PRIOR APPLICATION NUMBER: 09/516,448
;; PRIOR FILING DATE: 2000-02-29
;; NUMBER OF SEQ ID NOS: 2050
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1334
;; LENGTH: 492
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(492)
;; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1334

Query Match 3.0%; Score 49; DB 9; Length 492;
Best Local Similarity 100.0%; Pred. No. 4.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1183 GAGGAGTTCGAGAACACACTTTCCAAAAGCAGCGAGGTATTCACCAT 1231
Db 68 GAGGAGTTCGAGAACACACTTTCCAAAAGCAGCGAGGTATTCACCAT 116

RESULT 30
US-09-876-143-1454
;; Sequence 1454, Application US/09876143
;; Publication No. US20040081958A1
;; GENERAL INFORMATION:
;; APPLICANT: Infigen Inc.
;; APPLICANT: EILERTSEN, KENNETH J.
;; APPLICANT: PFISTER-GENSKOW, MARTHA
;; APPLICANT: CHILDS, LYNETTE
;; APPLICANT: FORSYTHE, TODD
;; APPLICANT: BISHOP, MICHAEL D.
;; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
;; TITLE OF INVENTION: CELLULAR REPROGRAMMING
;; FILE REFERENCE: 028040-0202
;; CURRENT APPLICATION NUMBER: US/09/876,143
;; CURRENT FILING DATE: 2001-06-06
;; PRIOR APPLICATION NUMBER: 60/209,874
;; PRIOR FILING DATE: 2000-06-07
;; NUMBER OF SEQ ID NOS: 1744
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 1454
;; LENGTH: 937
;; TYPE: DNA
;; ORGANISM: Bovine
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(937)
;; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1454

Query Match 2.3%; Score 37; DB 12; Length 937;
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 313 GCTGAACCCGAGATGAGAGAGTCCCGGACCTATG 349
Db 198 GCTGAACCCGAGATGAGAGAGTCCCGGACCTATG 234

RESULT 31
US-10-425-114-6003
; Sequence 6003, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6003
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700550751_FLI
US-10-425-114-6003
Query Match 1.4%; Score 23; DB 13; Length 1351;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1233 CAAGCAGGAGATGGAAGATGA 1255
DB 669 CAAGCAGGAGATGGAAGATGA 691
RESULT 32
US-10-425-114-32431
; Sequence 32431, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32431
; LENGTH: 1955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73326B04_FLI
US-10-425-114-32431
Query Match 1.4%; Score 23; DB 13; Length 1955;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1233 CAAGCAGGAGATGGAAGATGA 1255
DB 1237 CAAGCAGGAGATGGAAGATGA 1259
RESULT 33
US-09-962-055-42
; Sequence 42, Application US/09962055

; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...21
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-09-962-055-42
Query Match 1.3%; Score 21; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 CGTGATGCTCTGAGGAGCTG 225
DB 1 CGTGATGCTCTGAGGAGCTG 21
RESULT 34
US-09-976-740-42
; Sequence 42, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12

;; PRIOR APPLICATION NUMBER: 09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 42
;; LENGTH: 21
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-976-740-42

Query Match 1.3%; Score 21; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGCTCTCTGAGGAGCTG 225
Db 1 CGTGATGCTCTCTGAGGAGCTG 21

RESULT 35
US-10-671-242-42
;; Sequence 42, Application US/10671242
;; Publication No. US20040040049A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING AND TREATING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/671,242
;; CURRENT FILING DATE: 2003-09-24
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 42
;; LENGTH: 21
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-671-242-42

Query Match 1.3%; Score 21; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGCTCTCTGAGGAGCTG 225
Db 1 CGTGATGCTCTCTGAGGAGCTG 21

RESULT 36
US-10-023-529-42
;; Sequence 42, Application US/10023529
;; Publication No. US20020129388A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.

;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING AND TREATING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,529
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: 09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 42
;; LENGTH: 21
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-023-529-42

Query Match 1.3%; Score 21; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CGTGATGCTCTCTGAGGAGCTG 225
Db 1 CGTGATGCTCTCTGAGGAGCTG 21

RESULT 37
US-10-023-523-42
;; Sequence 42, Application US/10023523
;; Publication No. US20020152485A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING AND TREATING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,523
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 42
;; LENGTH: 21
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-023-523-42

Query Match 1.3%; Score 21; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	205	CGTGATGTCTCTGAGGAGCTG	225
D _b	1	CGTGATGTCTCTGAGGAGCTG	21

RESULT, T 38

```

US-10-616-187-42
; Sequence 42, Application US/10616187
; Publication NO. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Aribal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-616-187-42

```

```
Query Match      1.3%;   Score 21;   DB 16;   Length 21;
Best Local Similarity 100.0%;   Pred.No. 8.9;
Matches 21; Conservative 0; Mismatches 0; Indels
```

QY 205 CGTGATGTCCTCTGAGGAGCTG 225
|||
Db 1 CGTGATGTCCTCTGAGGAGCTG 21

RESULT 39

```

US-10-125-968-417
; Sequence 417, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: 24, 113, 121
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-417

```

```

Query Match      1.3%; Score 21; DB 16; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 21; Conservative 0; Mismatches 0; Indels

```

Qy 127 GCAGAAAGGTCCCGGCAGCAGC 147
Db 39 GCAGAAAGGTCCCGGCAGCAGC 59

RESULT 40

```

US-10-085-783A-546/c
; Sequence 546, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 546
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-546

```

Query Match 1.3%; Score 21; DB 13; Length 160;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 21; Conservative 0; Mismatches 0; Indels

QY
127 GCAGAGGTCCCGCAGCAGC 147

Dβ
78 GCAGAGGTCCCGCAGCAGC 58

RESULT 41

US-242-535A-546/c
; Sequence 546, Application US/10242535A
; Publication NO. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 546
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Human

US-10-242-535A-546

Query Match 1.3%; Score 21; DB 16; Length 160;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

Db 78 GCAGAGGTCCCGCAGCAGC 58

RESULT 42

US-10-085-783A-11558/c
; Sequence 11558, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 11558
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (149)..(149)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-11558

Query Match 1.3%; Score 21; DB 13; Length 261;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

Db 107 GCAGAGGTCCCGCAGCAGC 87

RESULT 43

US-10-242-535A-11558/c
; Sequence 11558, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 11558
; LENGTH: 261
; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (149)..(149)

; OTHER INFORMATION: n is a, c, g, or t

US-10-242-535A-11558

Query Match 1.3%; Score 21; DB 16; Length 261;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

Db 107 GCAGAGGTCCCGCAGCAGC 87

RESULT 44

US-10-085-783A-3476/c
; Sequence 3476, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3476
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-3476

Query Match 1.3%; Score 21; DB 13; Length 340;

Best Local Similarity 100.0%; Pred. No. 7.1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147

Db 70 GCAGAGGTCCCGCAGCAGC 50

RESULT 45

US-10-242-535A-3476/c
; Sequence 3476, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3476
; LENGTH: 340

TYPE: DNA
ORGANISM: Human
US-10-242-535A-3476

Query Match 1.3%; Score 21; DB 16; Length 340;
Best Local Similarity 100.0%; Pred. No. 7.1; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

Qy 127 GCAGAAGGTCCTCCGCGCAGC 147
Db 70 GCAGAAGGTCCTCCGCGCAGC 50

RESULT 46

US-09-920-300A-1581/c
Sequence 1581, Application US/09920300A
Publication No. US20020136728A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.547
CURRENT APPLICATION NUMBER: US/09/920,300A
CURRENT FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 1789
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1581
LENGTH: 393
TYPE: DNA
ORGANISM: Homo sapiens
US-09-920-300A-1581

Query Match 1.3%; Score 21; DB 9; Length 393;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAAGGTCCTCCGCGCAGC 147
Db 90 GCAGAAGGTCCTCCGCGCAGC 70

RESULT 47

US-10-033-528-1581/c
Sequence 1581, Application US/10033528
Publication No. US20020131971A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.547C1
CURRENT APPLICATION NUMBER: US/10/033,528
CURRENT FILING DATE: 2001-12-26
NUMBER OF SEQ ID NOS: 1896
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1581
LENGTH: 393
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-528-1581

Query Match 1.3%; Score 21; DB 14; Length 393;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAAGGTCCTCCGCGCAGC 147
Db 90 GCAGAAGGTCCTCCGCGCAGC 70

RESULT 48

US-10-099-926-1581/c
Sequence 1581, Application US/10099926
Publication No. US20030166064A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
APPLICANT: Jiang, Yugu
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.547C2
CURRENT APPLICATION NUMBER: US/10/099,926
CURRENT FILING DATE: 2002-03-17
NUMBER OF SEQ ID NOS: 1982
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1581
LENGTH: 393
TYPE: DNA
ORGANISM: Homo sapiens
US-10-099-926-1581

Query Match 1.3%; Score 21; DB 15; Length 393;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAAGGTCCTCCGCGCAGC 147
Db 90 GCAGAAGGTCCTCCGCGCAGC 70

RESULT 49

US-10-085-783A-9227/c
Sequence 9227, Application US/10085783A
Publication No. US20040037841A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2002
CURRENT APPLICATION NUMBER: US/10/085,783A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9227
LENGTH: 400
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (7)..(7)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (14)..(14)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (44)..(44)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: n is a, c, g, or t

Qy 127 GCAGAAGGTCCTCCGCGCAGC 147
Db 90 GCAGAAGGTCCTCCGCGCAGC 70

FEATURE:
NAME/KEY: misc feature
LOCATION: (352)..(352)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (397)..(397)
OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-9227

Query Match 1.3%; Score 21; DB 13; Length 400;
Best Local Similarity 100.0%; Pred.No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 129 GCAGAGGTCCTCCGCGCAGC 109

RESULT 50
US-10-242-535A-9227/c
Sequence 9227, Application US/1024535A
Publication No. US20040013663A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
CURRENT APPLICATION NUMBER: US/10/242,535A
CURRENT FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 10/085,783
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9227
LENGTH: 400
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (7)..(7)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (14)..(14)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (44)..(44)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (352)..(352)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (397)..(397)
OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-9227

Query Match 1.3%; Score 21; DB 16; Length 400;
Best Local Similarity 100.0%; Pred.No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 129 GCAGAGGTCCTCCGCGCAGC 109

RESULT 51
US-10-198-846-6336
Sequence 6336, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6336
LENGTH: 403
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 293, 318, 371
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-6336

Query Match 1.3%; Score 21; DB 15; Length 403;
Best Local Similarity 100.0%; Pred.No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 45 GCAGAGGTCCTCCGCGCAGC 65

RESULT 52
US-09-814-353-2628
Sequence 2628, Application US/09814353
Publication No. US20030165831A1
GENERAL INFORMATION:
APPLICANT: Lee, John
APPLICANT: Thompson, Pamela
APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2628


```
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 93..399
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-2628

Query Match      1.3%; Score 21; DB 10; Length 412;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGAGCAGC 147
Db 36 GCAGAGGTCCTCCGCGAGCAGC 56

RESULT 53
US-09-814-353-8962
; Sequence 8962, Application US/09814353
; Publication No. US2003016583A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-25
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-25
; PRIOR FILING DATE: 2000-07-25
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8962
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 93..399
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-8962

Query Match      1.3%; Score 21; DB 10; Length 412;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGAGCAGC 147
Db 36 GCAGAGGTCCTCCGCGAGCAGC 56

RESULT 54
US-10-198-846-9365/c
; Sequence 9365, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
```

```
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9365
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7..11, 18, 28, 324, 325, 326, 327, 328, 329, 330, 331, 332,
; LOCATION: 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344,
; LOCATION: 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356,
; LOCATION: 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 368..369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379,
; LOCATION: 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391,
; LOCATION: 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403,
; LOCATION: 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 415..416
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9365

Query Match      1.3%; Score 21; DB 15; Length 416;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCTCCGCGAGCAGC 147
Db 162 GCAGAGGTCCTCCGCGAGCAGC 142

RESULT 55
US-09-814-353-14275
; Sequence 14275, Application US/09814353
; Publication No. US2003016583A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14275
```

```
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-14275

Query Match      1.3%; Score 21; DB 10; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAAGTCCCGCAGCAGC 147
      |||||||
Db      34 GCAGAAGTCCCGCAGCAGC 54

RESULT 56
US-09-777-564-423/c
; Sequence 423; Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 423
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-777-564-423

Query Match      1.3%; Score 21; DB 9; Length 446;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAAGTCCCGCAGCAGC 147
      |||||||
Db      68 GCAGAAGTCCCGCAGCAGC 48

RESULT 57
US-10-015-219-423/c
; Sequence 423; Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015,219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 423
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-219-423

Query Match      1.3%; Score 21; DB 15; Length 446;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAAGTCCCGCAGCAGC 147
      |||||||
Db      68 GCAGAAGTCCCGCAGCAGC 48

RESULT 58
US-09-777-564-139/c
; Sequence 139; Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 139
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(447)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-564-139

Query Match      1.3%; Score 21; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAAGTCCCGCAGCAGC 147
      |||||||
Db      68 GCAGAAGTCCCGCAGCAGC 48

RESULT 59
US-10-015-219-139/c
; Sequence 139; Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015,219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 139
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 313, 412
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-139

Query Match      1.3%; Score 21; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAAGTCCCGCAGCAGC 147
      |||||||
Db      68 GCAGAAGTCCCGCAGCAGC 48

RESULT 60
US-09-920-455-171/c
; Sequence 171; Application US/09920455
; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
```

```
; CURRENT APPLICATION NUMBER: US/09/920,455
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 171
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 5, 6, 21, 405, 441
; OTHER INFORMATION: n = A,T,C or G
US-09-920-455-171

Query Match
Best Local Similarity 100.0%; Pred. No. 6.9; Length 458;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 99 GCAGAGGTCCTCCGCGCAGC 79

RESULT 61
US-09-814-353-14284/c
; Sequence 14284, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14284
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 26
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-14284
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 6.9; Length 462;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 229 GCAGAGGTCCTCCGCGCAGC 209
```

```
RESULT 62
US-09-814-353-15346
; Sequence 15346, Application US/09814353
```

```
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15346
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-15346

Query Match
Best Local Similarity 100.0%; Pred. No. 6.9; Length 497;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGCGCAGC 147
Db 34 GCAGAGGTCCTCCGCGCAGC 54

RESULT 63
US-09-895-828-75/c
; Sequence 75, Application US/09895828
; Patent No. US2002009012A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Carter, Darick
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.539
; CURRENT APPLICATION NUMBER: US/09/895,828
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 172_358, 400, 422
; OTHER INFORMATION: n = A,T,C or G
US-09-895-828-75

Query Match
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 127 GCAGAGGTCCTCCGCGCAGC 147
```

```
Db 95 GCAGAGGTCCCGCAGCAGC 75
|||||
RESULT 64
US-10-114-666-75/c
; Sequence 75, Application US/10114666
; Publication No. US20030103994A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kallos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.539C1
; CURRENT APPLICATION NUMBER: US/10/114,666
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 172, 358, 400, 422
; OTHER INFORMATION: n = A,T,C or G
US-10-114-666-75
Query Match 1.3%; Score 21; DB 15; Length 506;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCCGCAGCAGC 147
|||||
Db 95 GCAGAGGTCCCGCAGCAGC 75

RESULT 65
US-09-918-995-31626/c
; Sequence 31626, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31626
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(538)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31626
Query Match 1.3%; Score 21; DB 10; Length 538;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCCGCAGCAGC 147
Db 132 GCAGAGGTCCCGCAGCAGC 112

RESULT 66
```

```
US-09-813-358-167/c
; Sequence 167, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813,358
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(608)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-358-167
Query Match 1.3%; Score 21; DB 9; Length 608;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCCGCAGCAGC 147
Db 97 GCAGAGGTCCCGCAGCAGC 77

RESULT 67
US-09-997-279-167/c
; Sequence 167, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997,279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(608)
; OTHER INFORMATION: n = A,T,C or G
US-09-997-279-167
Query Match 1.3%; Score 21; DB 10; Length 608;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAGGTCCCGCAGCAGC 147
Db 97 GCAGAGGTCCCGCAGCAGC 77

RESULT 68
US-10-002-631C-81
; Sequence 81, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
```

; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-002-631C-81

Query Match 1.3%; Score 21; DB 15; Length 631;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
Db 536 GCAGAGGTCCCGCAGCAGC 556

RESULT 69

US-09-822-849A-69/c
; Sequence 69, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecthel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-69

Query Match 1.3%; Score 21; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
Db 97 GCAGAGGTCCCGCAGCAGC 77

RESULT 70

US-09-822-849A-68/c
; Sequence 68, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Fecthel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.

; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-68

Query Match 1.3%; Score 21; DB 9; Length 857;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
Db 97 GCAGAGGTCCCGCAGCAGC 77

RESULT 71

US-09-880-107-2377/c
; Sequence 2377, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2377
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M63573
US-09-880-107-2377

Query Match 1.3%; Score 21; DB 9; Length 893;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCAGC 147
Db 116 GCAGAGGTCCCGCAGCAGC 96

RESULT 72

US-09-960-706-697/c
; Sequence 697, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07

```
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 697
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M63573
US-09-960-706-697

Query Match      1.3%; Score 21; DB 10; Length 893;
Best Local Similarity 100.0%; Pred.No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAGGTCCTCCGGCAGCAGC 147
      |||||||
Db      116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 73
US-09-873-319-439/c
; Sequence 439, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Mueger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 439
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M63573
US-09-873-319-439

Query Match      1.3%; Score 21; DB 10; Length 893;
Best Local Similarity 100.0%; Pred.No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAGGTCCTCCGGCAGCAGC 147
      |||||||
Db      116 GCAGAGGTCCTCCGGCAGCAGC 96

RESULT 74
US-09-925-301-481/c
; Sequence 481, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-05
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 481
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-925-301-481

Query Match      1.3%; Score 21; DB 9; Length 1119;
Best Local Similarity 100.0%; Pred.No. 6.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAGGTCCTCCGGCAGCAGC 147
      |||||||
Db      320 GCAGAGGTCCTCCGGCAGCAGC 300

RESULT 75
US-10-264-049-799/c
; Sequence 799, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 799
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119)..(1120)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-799

Query Match      1.3%; Score 21; DB 16; Length 1131;
Best Local Similarity 100.0%; Pred.No. 6.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 GCAGAGGTCCTCCGGCAGCAGC 147
      |||||||
Db      320 GCAGAGGTCCTCCGGCAGCAGC 300

RESULT 76
US-10-198-846-13193/c
; Sequence 13193, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13193
; LENGTH: 1286
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-10-198-846-13193

Query Match 1.3%; Score 21; DB 15; Length 1286;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCTCCGGCAGCAGC 147
DB 190 GCAGAGGTCCTCCGGCAGCAGC 170

RESULT 77

US-09-960-352-10281/c
; Sequence 10281, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; TITLE OF INVENTION: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10281
; LENGTH: 328
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 44-LIB34-004-Q1-E1-C8
US-09-960-352-10281

Query Match 1.2%; Score 20; DB 9; Length 328;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 GCGCCGGGAGGAGGAGGAGA 805
DB 239 GCGCCGGGAGGAGGAGGAGA 220

RESULT 78

US-10-027-632-290731/c
; Sequence 290731, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290731
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-290731

Query Match 1.2%; Score 20; DB 13; Length 570;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1169 ACACAGAGAAGTTTGAGGAG 1188
DB 32 ACACAGAGAAGTTTGAGGAG 13

RESULT 79

US-10-027-632-290731/c
; Sequence 290731, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290731
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-290731

Query Match 1.2%; Score 20; DB 16; Length 570;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1169 ACACAGAGAAGTTTGAGGAG 1188
DB 32 ACACAGAGAAGTTTGAGGAG 13

RESULT 80

US-09-938-842A-1855/c
; Sequence 1855, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1855
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1855

Query Match 1.2%; Score 20; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CGGAGGAGGAGGAGCG 809
Db 289 CGGAGGAGGAGGAGCG 270

RESULT 81

US-09-938-842A-1855/c
; Sequence 1855, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1855
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1855

Query Match 1.2%; Score 20; DB 11; Length 576;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CGGAGGAGGAGGAGCG 809
Db 289 CGGAGGAGGAGGAGCG 270

RESULT 82

US-09-960-253-172
; Sequence 172, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Radoch
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-172

Query Match 1.2%; Score 20; DB 9; Length 2045;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCAGAGGAGGAGCTGGC 569
Db 502 ACCCAGAGGAGGAGCTGGC 521

RESULT 83

US-10-115-831-34
; Sequence 34, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 2356
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)..(1554)
US-10-115-831-34

Query Match 1.2%; Score 20; DB 16; Length 2356;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 ACCCAGAGGAGGAGCTGGC 569
Db 418 ACCCAGAGGAGGAGCTGGC 437

RESULT 84

US-09-764-847-1624
; Sequence 1624, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1624
; LENGTH: 20029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1624

Query Match 1.2%; Score 20; DB 9; Length 20029;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1170 CACAGAGAAGTTTGAGGAGT 1189


```
Db 15291 CACAGAGAAGTTTGAGGAGT 15310

RESULT 85
US-10-092-154-1624
; Sequence 1624, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1624
; LENGTH: 20029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1624

Query Match 1.2%; Score 20; DB 15; Length 20029;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1170 CACAGAGAAGTTTGAGGAGT 1189
Db 15291 CACAGAGAAGTTTGAGGAGT 15310

RESULT 86
US-10-085-117-100
; Sequence 100, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 96599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-100

Query Match 1.2%; Score 20; DB 16; Length 96599;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1235 AGCAGAGATGGAAGATG 1254
Db 59264 AGCAGAGATGGAAGATG 59283

RESULT 87
US-10-125-968-249/c
; Sequence 249, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

Db 15291 CACAGAGAAGTTTGAGGAGT 15310

TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-032
CURRENT APPLICATION NUMBER: US/10/125,968
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/285,163
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 1417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 249
LENGTH: 138
TYPE: DNA
ORGANISM: Homo sapiens
US-10-125-968-249

Query Match 1.2%; Score 19; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 AGAAGGTCCCGCAGCAGC 147
Db 138 AGAAGGTCCCGCAGCAGC 120

RESULT 88
US-09-814-353-1541/c
; Sequence 1541, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1541
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16, 85, 189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-1541

Query Match 1.2%; Score 19; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 AGAAGGTCCCGCAGCAGC 147
Db 217 AGAAGGTCCCGCAGCAGC 199

RESULT 89
US-09-814-353-7900/c
```

```
; Sequence 7900, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7900
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16, 85, 189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-7900

Query Match 1.2% Score 19; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 AGAAGTCCCGGCGAGC 147
| | | | | | | | | |
Db 217 AGAAGTCCCGGCGAGC 199

RESULT 90
US-09-814-353-1532
; Sequence 1532, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7900
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16, 85, 189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-7900

Query Match 1.2% Score 19; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 AGAAGTCCCGGCGAGC 147
| | | | | | | | | |
Db 217 AGAAGTCCCGGCGAGC 199

RESULT 90
US-09-814-353-1532
; Sequence 1532, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7900
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16, 85, 189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-7900
```

```
; SEQ ID NO 1532
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-1532

Query Match 1.2% Score 19; DB 10; Length 230;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAAGTCCCGGCGAGCA 145
| | | | | | | | | |
Db 34 GCAGAAGTCCCGGCGAGCA 52

RESULT 91
US-09-814-353-7891
; Sequence 7891, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7891
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-7891

Query Match 1.2% Score 19; DB 10; Length 230;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCAGAAGTCCCGGCGAGCA 145
| | | | | | | | | |
Db 34 GCAGAAGTCCCGGCGAGCA 52

RESULT 92
US-10-085-783A-40305
; Sequence 40305, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
```

; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40305
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-40305

Query Match 1.2%; Score 19; DB 13; Length 472;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1532 AAGCACCAGCTCCCCAG 1550
DB 148 AAGCACCAGCTCCCCAG 166

RESULT 93

US-10-242-535A-40305
; Sequence 40305, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40305
; LENGTH: 472

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)

; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-40305

Query Match 1.2%; Score 19; DB 16; Length 472;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1532 AAGCACCAGCTCCCCAG 1550
DB 148 AAGCACCAGCTCCCCAG 166

RESULT 94

US-10-198-846-9413
; Sequence 9413, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:

; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9413
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5, 6, 10, 77, 279, 372, 439, 440, 442, 469, 472, 479, 486,
; LOCATION: 494, 498, 511, 519
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9413

Query Match 1.2%; Score 19; DB 15; Length 519;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCAGAGGTCCCGCAGCA 145
DB 58 GCAGAGGTCCCGCAGCA 76

RESULT 95

US-10-027-632-181095
; Sequence 181095, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181095

; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-181095

Query Match 1.2%; Score 19; DB 13; Length 522;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1618 GAGCCACCTCCGCAGGG 1636
DB 70 GAGCCACCTCCGCAGGG 88

RESULT 96

US-10-027-632-317786

Sequence 317786, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317786
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317786

Query Match 1.2%; Score 19; DB 13; Length 522;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1618 GAGCCACCTCCGCCAGG 1636
DB 70 GAGCCACCTCCGCCAGG 88

RESULT 97
US-10-027-632-181095
; Sequence 181095, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181095
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-181095

Query Match 1.2%; Score 19; DB 16; Length 522;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1618 GAGCCACCTCCGCCAGG 1636
DB 70 GAGCCACCTCCGCCAGG 88

RESULT 98
US-10-027-632-317786
; Sequence 317786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317786
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317786

Query Match 1.2%; Score 19; DB 16; Length 522;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1618 GAGCCACCTCCGCCAGG 1636
DB 70 GAGCCACCTCCGCCAGG 88

RESULT 99
US-10-027-632-285491
; Sequence 285491, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

;
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285491
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-285491

Query Match 1.2%; Score 19; DB 13; Length 541;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1510 AGGCCAGAGGGCGCTGGG 1528
DB 407 AGGCCAGAGGGCGCTGGG 425

RESULT 100
US-10-027-632-285491
; Sequence 285491, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285491
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-285491

Query Match 1.2%; Score 19; DB 16; Length 541;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1510 AGGCCAGAGGGCGCTGGG 1528
DB 407 AGGCCAGAGGGCGCTGGG 425

Search completed: June 9, 2004, 01:38:23
Job time : 646.945 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 10:41:54 ; Search time 17858.2 Seconds
(without alignments)
11399.905 Million cell updates/sec

Title: US-10-023-523-17

Perfect score: 4697
Sequence: 1 caaaagcagccaggacaa.....ttaaaaaaaaaaaaaaaaaa 4697

Scoring table: OLIGO NUC
Gapop_60.0, Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database: GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_fod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4697	100.0	4697	6	AR199536	Sequence
2	4697	100.0	4697	6	AR374588	Sequence
3	4697	100.0	4697	6	AR409323	Sequence
4	4697	100.0	4697	6	AX239574	Sequence
5	4697	100.0	4697	6	BD056450	Novel low
6	1938	41.3	4880	9	HS803946	Homo sapi
7	1837	39.1	4851	9	HS803646	Homo sapi
8	1698	36.2	1593	9	AF516206	Homo sapi
9	1642	35.0	5023	9	HS803947	Homo sapi
10	1439	30.6	1638	6	AR409338	Sequence
11	1439	30.6	1638	6	AX239603	Sequence
12	1294	27.5	1793	6	AX578043	Sequence
13	1156	24.6	1793	9	HUNHMBGCF	Human high
14	1017	21.7	3554	9	BC046565	Homo sapi
15	921	19.6	3495	9	BC029686	Homo sapi
16	895	18.1	115756	9	HSDJ622L5	Human DNA
17	844	18.0	22255	6	AR409342	Sequence
18	844	18.0	22255	6	AX239608	Sequence
19	640	13.6	2713	9	HS802997	Homo sapi
20	635	13.5	3049	9	AK096161	Homo sapi
21	538	11.5	1477	9	AF070546	Homo sapi
22	367	7.8	529	6	BD060509	Secreted
23	346	7.4	4054	9	AK126027	Homo sapi
24	239	5.1	316	6	AX893880	Sequence
25	239	5.1	316	6	BD029413	Sequence
26	235	5.0	235	6	AX396823	Sequence
27	230	4.9	2194	6	AX834462	Sequence
28	230	4.9	2194	9	AK096939	Homo sapi
29	225	4.8	250	6	AX209535	Sequence
30	225	4.8	370	6	AX199010	Sequence
31	173	3.7	421	6	AR425015	Sequence
32	173	3.7	421	6	BD120568	EST and e
33	166	3.5	441	6	AX331983	Sequence
34	1474	3.1	386	11	G28385	human STS S
35	99	2.1	383	11	G06968	human STS W
36	59	1.3	4722	6	AR199533	Sequence
37	59	1.3	4722	6	AR374685	Sequence
38	59	1.3	4722	6	AR409320	Sequence
39	59	1.3	4722	6	AX239571	Sequence
40	59	1.3	4722	6	BD056447	Novel low
41	49	1.0	142	11	G26511	human STS T
42	49	1.0	245095	2	AC111774	Rattus no
43	47	1.0	4702	10	BC060227	Mus muscu
44	47	1.0	141790	10	AL671759	Mouse DNA
45	44	0.9	103829	9	AF000665	Homo sapi
46	44	0.9	178526	2	AF001192	Homo sapi
47	44	0.9	181547	2	AC108815	Mus muscu
48	40	0.9	7622	3	AF112367	Plasmodu
49	38	0.8	600	9	BC041996	Homo sapi
50	38	0.8	605	9	BC033959	Homo sapi
51	38	0.8	143563	9	H855C23	Human DNA
52	37	0.8	236171	2	AC105852	Rattus no
53	37	0.8	241633	2	AC096932	Rattus no
54	37	0.8	245826	2	AC099192	Rattus no
55	37	0.8	272182	2	AC111946	Rattus no
56	36	0.8	110000	2	PFMA6P1.05	Continuation (6 of
57	35	0.7	401	11	BV048056	S212P6400
58	35	0.7	162837	9	AL451137	Human DNA
59	35	0.7	165287	2	AC133880	Mus muscu
60	35	0.7	166952	2	AC073836	Homo sapi
61	35	0.7	194006	2	AC127441	Rattus no
62	35	0.7	226253	2	AC133434	Rattus no
63	35	0.7	227996	2	AC117916	Rattus no
64	35	0.7	263236	2	AC094070	Rattus no
65	35	0.7	270897	2	AC096415	Rattus no

```

c 66 34 0.7 14615 6 AX251462 Sequence
c 67 34 0.7 10000 2 AC073763 Mus muscu
c 68 34 0.7 130254 2 AC136195 Rattus no
c 69 34 0.7 13516 9 AC083783 Homo sapi
c 70 34 0.7 15439 9 AC073161 Homo sapi
c 71 34 0.7 164858 9 AC026634 Homo sapi
c 72 34 0.7 165316 2 AC119115 Rattus no
c 73 34 0.7 169539 9 AC021321 Homo sapi
c 74 34 0.7 183235 2 AC020714 Homo sapi
c 75 34 0.7 192326 2 AC134317 Rattus no
c 76 34 0.7 218087 10 AC113082 Mus muscu
c 77 34 0.7 220702 2 AC114181 Rattus no
c 78 34 0.7 222739 2 AC095465 Rattus no
c 79 34 0.7 223877 2 AC098923 Rattus no
c 80 34 0.7 225077 2 AC136563 Rattus no
c 81 34 0.7 225253 2 AC136482 Rattus no
c 82 34 0.7 230760 2 AC073714 Mus muscu
c 83 34 0.7 231117 2 AC108548 Rattus no
c 84 34 0.7 245993 2 AC130970 Rattus no
c 85 34 0.7 253343 10 AC102566 Mus muscu
c 86 34 0.7 254951 2 AC122999 Rattus no
c 87 34 0.7 255298 2 AC135454 Rattus no
c 88 34 0.7 260860 2 AC125982 Rattus no
c 89 34 0.7 288968 2 AC115432 Rattus no
c 90 33 0.7 4335 9 AC098792 Homo sapi
c 91 33 0.7 66707 2 AC090878 Homo sapi
c 92 33 0.7 110000 2 AC024907 Homo sapi
c 93 33 0.7 113457 9 AL357518 Human DNA
c 94 33 0.7 134457 9 AC001231 Homo sapi
c 95 33 0.7 139684 2 AC130483 Mus muscu
c 96 33 0.7 143773 9 AC005631 Homo sapi
c 97 33 0.7 154162 2 AC069270 Homo sapi
c 98 33 0.7 156495 2 AC127977 Rattus no
c 99 33 0.7 162273 2 AL135909 Homo sapi
c 100 33 0.7 165378 2 AC143332 Homo sapi

```

ALIGNMENTS

```

RESULT 1
LOCUS AR199536 4697 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 17 from patent US 6355451.
ACCESSION AR199536
VERSION AR199536.1 GI:20249610
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4697)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6355451-A 17 12-MAR-2002;
FEATURES
source Location/Qualifiers
1. 4697
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 4697; DB 6; Length 4697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAAAGCAGCCAGGACCAACCGAAGCAGGAGCCGAGGAGCCAGAGCGCCAGGCC 60
DB 1 CAAAAGCAGCCAGGACCAACCGAAGCAGGAGCCGAGGAGCCAGAGCGCCAGGCC 60
QY 61 AGGCGGCTCTCGAGTAGAAGCAAGAGTCCCGGCAGCAGCCAGGCTCTCGGAAGCCGG 120
DB 61 AGGCGGCTCTCGAGTAGAAGCAAGAGTCCCGGCAGCAGCCAGGCTCTCGGAAGCCGG 120

```

```

121 AGGGGGCTCAAGCCAGAACGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 180
121 AGGGGGCTCAAGCCAGAACGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 180
181 GCGGCCAATCGAGACATCTGTGACATATCTGTGTGCAATAAACCAGAGGGGCCCCG 240
181 GCGGCCAATCGAGACATCTGTGACATATCTGTGTGCAATAAACCAGAGGGGCCCCG 240
241 GCGAGATGGGGCACAAGGCTGAGCGGCTGAACCCGGAAGATGCAGAGAAGTCCCGGACCT 300
241 GCGAGATGGGGCACAAGGCTGAGCGGCTGAACCCGGAAGATGCAGAGAAGTCCCGGACCT 300
301 ATGTGGCAAGGATGGGAGCTGAACCACTCCAGTAGTCTATGGAGAGAGAACCTT 360
301 ATGTGGCAAGGATGGGAGCTGAACCACTCCAGTAGTCTATGGAGAGAGAACCTT 360
361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACC 420
361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACC 420
421 ATCGAAGGCCACAGAGAGAGAGAGAAAGCAAGGGTTTGGGGAAGAGATCATCGTTGCTGA 480
421 ATCGAAGGCCACAGAGAGAGAGAGAAAGCAAGGGTTTGGGGAAGAGATCATCGTTGCTGA 480
481 TGCAGACATTTGAATCTCTGAGTACCCAGAGGAGAGAGTGGCTGCTGTGCAAGAGT 540
481 TGCAGACATTTGAATCTCTGAGTACCCAGAGGAGAGAGTGGCTGCTGTGCAAGAGT 540
541 ATGCTGAATCTGTGAGGAGACACCGAATTCACAGAACAGATGAAGTCTCTACAGAAA 600
541 ATGCTGAATCTGTGAGGAGACACCGAATTCACAGAACAGATGAAGTCTCTACAGAAA 600
601 ACAGAGCAGCTGTGCAAGAGAGAGACACCTTCGGGCTGAGCAGAGAGAGCCGTC 660
601 ACAGAGCAGCTGTGCAAGAGAGAGAGACACCTTCGGGCTGAGCAGAGAGAGCCGTC 660
661 TGGCCCGCAGCAAGCTTGAGAGCCTATGCTGCTGAGTGCAGCGGCACAAACCGCTCCTCA 720
661 TGGCCCGCAGCAAGCTTGAGAGCCTATGCTGCTGAGTGCAGCGGCACAAACCGCTCCTCA 720
721 AGGAAGAGTGTGAGCGGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGC 780
721 AGGAAGAGTGTGAGCGGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGC 780
781 ACTTCCAGTGTACCTGAATGATTCAGTTCAGATGCAACAGACACATCAGAGGCACT 840
781 ACTTCCAGTGTACCTGAATGATTCAGTTCAGATGCAACAGACACATCAGAGGCACT 840
841 CCAAGCTGCGCCACAGAGAACATGAGAGTGGCTGAGAGCTCAAGAGCTGATTTGAGCAGT 900
841 CCAAGCTGCGCCACAGAGAACATGAGAGTGGCTGAGAGCTCAAGAGCTGATTTGAGCAGT 900
901 ATGAGCTGCGGAGAGAGATTCGCAAGAGTCTTCAACACAGAGAGCTCAACAGAGAGC 960
901 ATGAGCTGCGGAGAGAGATTCGCAAGAGTCTTCAACACAGAGAGCTCAACAGAGAGC 960
961 TGGTGGATGCCAAGCTCCAGAGGCCCCAGGAGATGCTTAAAGAGGAGAGAGAGAGAGC 1020
961 TGGTGGATGCCAAGCTCCAGAGGCCCCAGGAGATGCTTAAAGAGGAGAGAGAGAGAGC 1020
1021 AGCGGAGAGAGATTTCTCTGGAAGAGGAGAGTACCTCCAGAGGAGATGTGAGCTGA 1080
1021 AGCGGAGAGAGATTTCTCTGGAAGAGGAGAGTACCTCCAGAGGAGATGTGAGCTGA 1080
1081 TGAAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 1140
1081 TGAAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 1140
1141 AGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 1200
1141 AGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 1200
1201 AAAAGATGACTAAGAGATCAAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGTGG 1260

```


Db 1201 AAAAGATGCTAAGAGATCAAGAGCTGAGAGAAAGAACCCATGTACCGGTCCCGT 1260
Qy 1261 GGAGAGCAGCAACAAGGCCCTCTTGAAGTGGCTGAGGAGAAAACAGTCCGGGATAAAG 1320
Db 1261 GGAGAGCAGCAACAAGGCCCTCTTGAAGTGGCTGAGGAGAAAACAGTCCGGGATAAAG 1320
Qy 1321 AACTGAGGCGCTGAGAGTAAATCAACGGCTGGAGAAAGCTGTGCGGGCACTGCAGA 1380
Db 1321 AACTGAGGCGCTGAGAGTAAATCAACGGCTGGAGAAAGCTGTGCGGGCACTGCAGA 1380
Qy 1381 CAGAGCCTAATGACCTGAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCC 1440
Db 1381 CAGAGCCTAATGACCTGAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCC 1440
Qy 1441 TCACTGACGTGGCCCTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGACCCAGCTCC 1500
Db 1441 TCACTGACGTGGCCCTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGACCCAGCTCC 1500
Qy 1501 CCAGGGTACAGAGCGCTTGTACCCAGGAGCAGGAGCAGAGAGATCAGGGCAGA 1560
Db 1501 CCAGGGTACAGAGCGCTTGTACCCAGGAGCAGGAGCAGAGAGATCAGGGCAGA 1560
Qy 1561 CTGGGCTCAAGAGCCACCTCCGCGAGGCTAGAGAGCCTGGTGTGGTCACTGG 1620
Db 1561 CTGGGCTCAAGAGCCACCTCCGCGAGGCTAGAGAGCCTGGTGTGGTCACTGG 1620
Qy 1621 GAAGGAGCGGCGCCAGGAGCCTGGGCGCTAAAGGCTCCCATGCTGAGCAGCCCA 1680
Db 1621 GAAGGAGCGGCGCCAGGAGCCTGGGCGCTAAAGGCTCCCATGCTGAGCAGCCCA 1680
Qy 1681 TTGCTGAAGCAGGATGCTTCTGACCTGGCTGGCGCATCTGGCACTTGCAATTTTGGATTT 1740
Db 1681 TTGCTGAAGCAGGATGCTTCTGACCTGGCTGGCGCATCTGGCACTTGCAATTTTGGATTT 1740
Qy 1741 GTGGGTGAGTTTACGTACATAGGCAATTTTGAAGGCTTGAAGGCTTGAATTTATACCTG 1800
Db 1741 GTGGGTGAGTTTACGTACATAGGCAATTTTGAAGGCTTGAAGGCTTGAATTTATACCTG 1800
Qy 1801 TAAGTGTACAGTGGGCTTGAATTTGGGATGGGCTGTACAGATGAAGTCACTGGCTTG 1860
Db 1801 TAAGTGTACAGTGGGCTTGAATTTGGGATGGGCTGTACAGATGAAGTCACTGGCTTG 1860
Qy 1861 TCTGTAGCTGAAGAGTCTTGAAGGGGCTGTATCTGTAGCTGCCATCAGTGAAGTTG 1920
Db 1861 TCTGTAGCTGAAGAGTCTTGAAGGGGCTGTATCTGTAGCTGCCATCAGTGAAGTTG 1920
Qy 1921 GCAGAGTGTAGTGAAGTCTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCTTT 1980
Db 1921 GCAGAGTGTAGTGAAGTCTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCTTT 1980
Qy 1981 CAGAGCTCAAAACAAGTAATACACCAAGGCTTGTGCTGCAATTTGTGTGAGCAGGGCT 2040
Db 1981 CAGAGCTCAAAACAAGTAATACACCAAGGCTTGTGCTGCAATTTGTGTGAGCAGGGCT 2040
Qy 2041 TGCTTGTGCTAGCTCAGGCGCTCTAGTGTCTGAGGCTCCTTTGATCTCTAGAGCTGG 2100
Db 2041 TGCTTGTGCTAGCTCAGGCGCTCTAGTGTCTGAGGCTCCTTTGATCTCTAGAGCTGG 2100
Qy 2101 AAAAGTGTCCCTAGGAGGAGCCTGGCAGGGGCTCAGAGCTGGGATTTCTGCTGGA 2160
Db 2101 AAAAGTGTCCCTAGGAGGAGCCTGGCAGGGGCTCAGAGCTGGGATTTCTGCTGGA 2160
Qy 2161 ACAAGGACCTGGAGAAATTTTGGCTGGGATGATGTGTGAGGAGCCCTTTGGGC 2220
Db 2161 ACAAGGACCTGGAGAAATTTTGGCTGGGATGATGTGTGAGGAGCCCTTTGGGC 2220
Qy 2221 ATGCTTTCCCTGCTGGTGTGAGTGGCAGGAGCAGGCAATGATGCTTCTAGTAGCCT 2280
Db 2221 ATGCTTTCCCTGCTGGTGTGAGTGGCAGGAGCAGGCAATGATGCTTCTAGTAGCCT 2280
Qy 2281 TATCAITTCAGAGTGGCTCTCTAGCTGCAAAATGATTCAGAGAGATCACCCAAAGGA 2340

Db 2281 TATCAITTCAGAGTGGCTCTCTAGCTGCAAAATGATTCAGAGAGATCACCCAAAGGA 2340
Qy 2341 TTAITTCGAAGTGTCTTTTCTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2400
Db 2341 TTAITTCGAAGTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2400
Qy 2401 TTTTCTTCCATGACATGACAGTGTCTTGAATGAGGACCTTCCAAAGGAAAAGGATGCTGTACC 2460
Db 2401 TTTTCTTCCATGACATGACAGTGTCTTGAATGAGGACCTTCCAAAGGAAAAGGATGCTGTACC 2460
Qy 2461 AGTGTGCTCGGCTGGCTCCAGTGTCCACCTCTTCCACACCCCACTTGGCTCC 2520
Db 2461 AGTGTGCTCGGCTGGCTCCAGTGTCCACCTCTTCCACACCCCACTTGGCTCC 2520
Qy 2521 TTTGCACTCTGAGCTCAGGTTTCTGTTTGGTGGATCAGGTTGTTTGGTGGTAAAGA 2580
Db 2521 TTTGCACTCTGAGCTCAGGTTTCTGTTTGGTGGATCAGGTTGTTTGGTGGTAAAGA 2580
Qy 2581 AAGGAAAGGGCTTCTGATGGCTTGGCCACAAGCTTACCTGTGGGTTTCACTGCTGAGAG 2640
Db 2581 AAGGAAAGGGCTTCTGATGGCTTGGCCACAAGCTTACCTGTGGGTTTCACTGCTGAGAG 2640
Qy 2641 CCACCAAGTTCCTCATCAGCACTGTCTCAATGAGAGAGTGTCTGGTCCCATGCTCCAGC 2700
Db 2641 CCACCAAGTTCCTCATCAGCACTGTCTCAATGAGAGAGTGTCTGGTCCCATGCTCCAGC 2700
Qy 2701 TGCCTCTTGGCTTCTGATGGTCTTCTGCTTCCCTGCGCCCACTGCTGCAATCT 2760
Db 2701 TGCCTCTTGGCTTCTGATGGTCTTCTGCTTCCCTGCGCCCACTGCTGCAATCT 2760
Qy 2761 CAAGATTTGTCTGATTTCTATTTCTGCGACCTCTCTGCGCTTCTCTGGGATTTCTACT 2820
Db 2761 CAAGATTTGTCTGATTTCTATTTCTGCGACCTCTCTGCGCTTCTCTGGGATTTCTACT 2820
Qy 2821 CTTCTGTGTGGGCGCCATAGCTGTGTCTAACAGGTAAAGAAATGAATTTGATTTGA 2880
Db 2821 CTTCTGTGTGGGCGCCATAGCTGTGTCTAACAGGTAAAGAAATGAATTTGATTTGA 2880
Qy 2881 CTGGGCGCCAGAAAATCCATAAAATGGCTGCAGACAGTTGTTTCTGTGCTCTTCTACCC 2940
Db 2881 CTGGGCGCCAGAAAATCCATAAAATGGCTGCAGACAGTTGTTTCTGTGCTCTTCTACCC 2940
Qy 2941 CCACCTCAGTACATTAATCTATGTGTGTAGAGCCATTTCTATATGCTGAATTTCTG 3000
Db 2941 CCACCTCAGTACATTAATCTATGTGTGTAGAGCCATTTCTATATGCTGAATTTCTG 3000
Qy 3001 CTGTTGCAAACTTGGCAGGTATTAGCCAGTGTGTGTCGAAGCAGTTTTCGGGACAAAC 3060
Db 3001 CTGTTGCAAACTTGGCAGGTATTAGCCAGTGTGTGTCGAAGCAGTTTTCGGGACAAAC 3060
Qy 3061 AGAATGACTCAGACCAAGATGGATGGTGGTGGCTTTGCTTCTGCTGTTTCTTCT 3120
Db 3061 AGAATGACTCAGACCAAGATGGATGGTGGTGGCTTTGCTTCTGCTGTTTCTTCT 3120
Qy 3121 TGAATGTCTGATTTGCTGAGGCTCCCTTCACTTCCATCTAGCCCACTCTTTAGCC 3180
Db 3121 TGAATGTCTGATTTGCTGAGGCTCCCTTCACTTCCATCTAGCCCACTCTTTAGCC 3180
Qy 3181 CTTTACCTTAAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
Db 3181 CTTTACCTTAAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
Qy 3241 TGAATGAGACCAAGATGGCAATGAATGGCAGACATTTCTCTCTCTGCTGCGCCAGG 3300
Db 3241 TGAATGAGACCAAGATGGCAATGAATGGCAGACATTTCTCTCTCTGCTGCGCCAGG 3300
Qy 3301 CAGGTACCACTGATCTGATCAGAGTTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 3360
Db 3301 CAGGTACCACTGATCTGATCAGAGTTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 3360
Qy 3361 GGTGCGCTCAAGCAGCTGTGTGAGTGTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Db 3361 GGTGCGCTCAAGCAGCTGTGTGAGTGTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420

QY	3421	TGTCCTGTGAAGTGTGTTCCCATGACCTTTTCTTCCCTTTGAATCCCTCTTGCTGGAG	3480
DB	3421	TGTCCTGTGAAGTGTGTTCCCATGACCTTTTCTTCCCTTTGAATCCCTCTTGCTGGAG	3480
QY	3481	TAGTCCTTGGCTTCTTCTTGCTCCAGTAGGCTTTTCTTACCCAGCCCTTGTCGCCAGG	3540
DB	3481	TAGTCCTTGGCTTCTTCTTGCTCCAGTAGGCTTTTCTTACCCAGCCCTTGTCGCCAGG	3540
QY	3541	CTAAGCTGTTACAAAGAGCTGCCAACTCACAGAGTTTTGTCTAGGCGAGAGGTGTCAGGGA	3600
DB	3541	CTAAGCTGTTACAAAGAGCTGCCAACTCACAGAGTTTTGTCTAGGCGAGAGGTGTCAGGGA	3600
QY	3601	AGAGCGAGAGGTATGCACTTCCCTTTGAAGAGAGGGGAAAGCCCTPACAGTGGCCACACA	3660
DB	3601	AGAGCGAGAGGTATGCACTTCCCTTTGAAGAGAGGGGAAAGCCCTPACAGTGGCCACACA	3660
QY	3661	TAATTGGCTGACACACACTTCAGCTACCTCTTAATGCTCTGTGAGGAGACTGGAGCTGCTG	3720
DB	3661	TAATTGGCTGACACACTTCAGCTACCTCTTAATGCTCTGTGAGGAGACTGGAGCTGCTG	3720
QY	3721	GATCCCAAGTGTGGTGTAGGAGCCACAGTCAGCAGAGTGGCCCAAGCTGGGTTTCCCA	3780
DB	3721	GATCCCAAGTGTGGTGTAGGAGCCACAGTCAGCAGAGTGGCCCAAGCTGGGTTTCCCA	3780
QY	3781	GGTCAGGAATGTGGGCCCGCCAGGCAAGTGCAGCTTTGCTCACAGTCCATCCATGCTTA	3840
DB	3781	GGTCAGGAATGTGGGCCCGCCAGGCAAGTGCAGCTTTGCTCACAGTCCATCCATGCTTA	3840
QY	3841	GACCTTCAGGCGAGTCTGCAGATGAGGTTCCCTACCTTTTCTTCTCTTCAATTGACAAA	3900
DB	3841	GACCTTCAGGCGAGTCTGCAGATGAGGTTCCCTACCTTTTCTTCTCTTCAATTGACAAA	3900
QY	3901	TCAACCAATCACTACAGCTGCTGTCTTCTGCTTTTCCAAAGTAGCCAGGTCCTGGGCCA	3960
DB	3901	TCAACCAATCACTACAGCTGCTGTCTTCTGCTTTTCCAAAGTAGCCAGGTCCTGGGCCA	3960
QY	3961	GATSCAGGGAGGTGCCATCCATGAGTGAAGGCCAGTGCTTCTCCACCTGGTGGTCC	4020
DB	3961	GATSCAGGGAGGTGCCATCCATGAGTGAAGGCCAGTGCTTCTCCACCTGGTGGTCC	4020
QY	4021	CACACTTGTGACCCCTCAGTTTTTAGGCCCAAGATCTGTGTGGTTTCTTAGATTGCTAGC	4080
DB	4021	CACACTTGTGACCCCTCAGTTTTTAGGCCCAAGATCTGTGTGGTTTCTTAGATTGCTAGC	4080
QY	4081	TTTTTCTCCAGGGAGCCACAGCAGGTGAAGTCAAGAGCCGATGGCTCTCTTAATAGTAA	4140
DB	4081	TTTTTCTCCAGGGAGCCACAGCAGGTGAAGTCAAGAGCCGATGGCTCTCTTAATAGTAA	4140
QY	4141	ATTGTTTTTCAGGGCCTTGTCCAGCTGAGAGCTTCATGTCCACCAAGATTCTGAGAGGTGC	4200
DB	4141	ATTGTTTTTCAGGGCCTTGTCCAGCTGAGAGCTTCATGTCCACCAAGATTCTGAGAGGTGC	4200
QY	4201	AGCAGCACTTTTTTTTTTATTTTGTGTTGTTTCCATGAGGTTATCGGACCATGGGCT	4260
DB	4201	AGCAGCACTTTTTTTTTTATTTTGTGTTGTTTCCATGAGGTTATCGGACCATGGGCT	4260
QY	4261	GAGCTCAGGCACTTTTCTGTAGGAGACTGTTATTTCTGTAAGATGGTTATTTAAACCTCC	4320
DB	4261	GAGCTCAGGCACTTTTCTGTAGGAGACTGTTATTTCTGTAAGATGGTTATTTAAACCTCC	4320
QY	4321	TCCACCCCATCAAGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCTGTGTGC	4380
DB	4321	TCCACCCCATCAAGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCTGTGTGC	4380
QY	4381	CAGGGGGAGGGCCCAAGGCTGTGAGCTGATTTCTCCAGCTGTGTGCCCCAGGCTTTCCGC	4440
DB	4381	CAGGGGGAGGGCCCAAGGCTGTGAGCTGATTTCTCCAGCTGTGTGCCCCAGGCTTTCCGC	4440
QY	4441	CTTGCAAGCACAGAGGTGTTCACCCCGAGGGACAGCCAGGCACTGTCTCTTTGGCCCTT	4500
DB	4441	CTTGCAAGCACAGAGGTGTTCACCCCGAGGGACAGCCAGGCACTGTCTCTCTTTGGCCCTT	4500

Qy	4501	CCTGGGGAAAGAGAGTGCCTTCCTGTCCTCGTAACATGCTTTCCATTATGAGCCCAACCCGGC	4561
Db	4501	CCTGGGGAAAGAGAGTGCCTTCCTGTCCTCGTAACATGCTTTCCATTATGAGCCCAACCCGGC	4561
Qy	4561	CACTCAGAGCTGTGTTGAAGCTGCACGTGGAGCTTTTTTGTCCTCTTTGGGTATTACAAC	4620
Db	4561	CACTCAGAGCTGTGTTGAAGCTGCACGTGGAGCTTTTTTGTCCTCTTTGGGTATTACAAC	4620
Qy	4621	AGCCAGGAGCTTCATTTTGATGTATTTTAAACCACCATTAATAAAGAGTCTGTGGCTTA	4680
Db	4621	AGCCAGGAGCTTCATTTTGATGTATTTTAAACCACCATTAATAAAGAGTCTGTGGCTTA	4680
Qy	4681	AAAAAAAAAAAAAAAAAAAA 4697	
Db	4681	AAAAAAAAAAAAAAAAAAAA 4697	
 RESULT 2			
LOCUS AR374688 4697 bp DNA linear PAT 18-DEC-2003			
DEFINITION Sequence 17 from patent US 6605588.			
ACCESSION AR374688			
VERSION AR374688.1 GI:40077503			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 4697)			
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.			
TITLE Low density lipoprotein binding proteins and their use in diagnosing and treating atherosclerosis			
JOURNAL Patent: US 6605588-A 17 12-AUG-2003;			
FEATURES			
source Location/Qualifiers			
1..4697			
/organism="unknown"			
/mol_type="genomic DNA"			
 ORIGIN			
Query Match 100.0%; Score 4697; DB 6; Length 4697;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CAAAAAGCAGCCAGACAAACCGAAGCAGGACCCGAGGAGGCCAGGAGCGGCCAGGCC	60
Db	1	CAAAAAGCAGCCAGACAAACCGAAGCAGGACCCGAGGAGGCCAGGAGCGGCCAGGCC	60
Qy	61	AGCGGCTCTGCAGTAGACAGAGAGTCCCGGACGACGACGCTCTCGGAAGCCGG	120
Db	61	AGCGGCTCTGCAGTAGACAGAGAGTCCCGGACGACGACGCTCTCGGAAGCCGG	120
Qy	121	AGGGGCTCAAAGCAGAACGGCTCAGTCTGGGCCCCCTTCGTGATGTCTCTGAGGAGCTGA	180
Db	121	AGGGGCTCAAGCCAGACAGGCTCAGTCTGGGCCCCCTTCGTGATGTCTCTGAGGAGCTGA	180
Qy	181	GCGGCCAATGGAAGACATACCTAGCACATCTGTGTGACAAATACCGAGGGGGCCCCG	240
Db	181	GCGGCCAATGGAAGACATACCTAGCACATCTGTGTGACAAATACCGAGGGGGCCCCG	240
Qy	241	GCGAGGATGGGGCACAGGGTGAGCCGGCTGAACCCCAAGATGCAGAGAAGTCCCAGACCT	300
Db	241	GCGAGGATGGGGCACAGGGTGAGCCGGCTGAACCCCAAGATGCAGAGAAGTCCCAGACCT	300
Qy	301	ATGTGCAGAGGAATGGGAGCCTGAACCAACTCCAGTAGTCTATGGAGAGGAACCCCT	360
Db	301	ATGTGCAGAGGAATGGGAGCCTGAACCAACTCCAGTAGTCTATGGAGAGGAACCCCT	360
Qy	361	CCAAGGGGATCCAAACACAGAAAGATCCGGCAGAGTGACAGGTCGGAGACCGAGACC	420
Db	361	CCAAGGGGATCCAAACACAGAAAGATCCGGCAGAGTGACAGGTCGGAGACCGAGACC	420
Qy	421	ATCGAAGGCCACAGGAGAGAAAAAACCAAGGGTTTGGGGAAGGAGATCACCTTGTCTGA	480
Db	421	ATCGAAGGCCACAGGAGAGAAAAAACCAAGGGTTTGGGGAAGGAGATCACCTTGTCTGA	480

QY 481 TGCGACATTTGAATCTCTGAGTACCCAGAGGAGAGCTGGCTGCTCTGTGTCAGAAAGT 540
Db 481 TGCGACATTTGAATCTCTGAGTACCCAGAGGAGAGCTGGCTGCTCTGTGTCAGAAAGT 540
QY 541 ATGCTGAATCTGTGGAGGACACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 600
Db 541 ATGCTGAATCTGTGGAGGACACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 600
QY 601 AGCAGAGCAGCTGTGTCAAGAGAGGACCACTGCGCGGTGAGCAGACAGAGGCGCTCC 660
Db 601 AGCAGAGCAGCTGTGTCAAGAGAGGACCACTGCGCGGTGAGCAGACAGAGGCGCTCC 660
QY 661 TGCGCCGACGAGCTTGAGAGCTATGCCGTGAGCTGAGCGGACAAACCGCTCCCTCA 720
Db 661 TGCGCCGACGAGCTTGAGAGCTATGCCGTGAGCTGAGCGGACAAACCGCTCCCTCA 720
QY 721 AGAAGAGGTGTGACGCGGCGCGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 AGAAGAGGTGTGACGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 ACTTCCAGGTGACACTGAAATGACATTCAGCTGCAGATGAAACAGCACAATGAGCGCACT 840
Db 781 ACTTCCAGGTGACACTGAAATGACATTCAGCTGCAGATGAAACAGCACAATGAGCGCACT 840
QY 841 CCAAGCTGCGCCAAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
Db 841 CCAAGCTGCGCCAAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
QY 901 ATGAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACAGAGGACCTACACAGCAGC 960
Db 901 ATGAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACAGAGGACCTACACAGCAGC 960
QY 961 TGGTGGATGCAAGCTCCAGCAGCGCCAGAGATGCTAAAGGAGGAGAGAGGAGCGGACCC 1020
Db 961 TGGTGGATGCAAGCTCCAGCAGCGCCAGAGATGCTAAAGGAGGAGAGAGGAGCGGACCC 1020
QY 1021 AGCGGAGAGAGATTTTCTCTCAAGAGGAGTACAGTCCAGAGATGCTGAGCTCA 1080
Db 1021 AGCGGAGAGAGATTTTCTCTCAAGAGGAGTACAGTCCAGAGATGCTGAGCTCA 1080
QY 1081 TGAAGCAGAGAGAGACCACTGAAGCAACAGCTTGCCTTATACACAGAGAGATTTGAGG 1140
Db 1081 TGAAGCAGAGAGAGACCACTGAAGCAACAGCTTGCCTTATACACAGAGAGATTTGAGG 1140
QY 1141 AGTTCCAGAACACACTTTCCAAAGCAGCGAGTATTCACCAATTCAGAGCAGGAGATGG 1200
Db 1141 AGTTCCAGAACACACTTTCCAAAGCAGCGAGTATTCACCAATTCAGAGCAGGAGATGG 1200
QY 1201 ABAAGATGACTAAGAGATCAAGAGCTGGAGAGAGAAACCACTGTATCCCGTCCCGGT 1260
Db 1201 ABAAGATGACTAAGAGATCAAGAGCTGGAGAGAGAAACCACTGTATCCCGTCCCGGT 1260
QY 1261 GGGAGAGCAGCAACAGGCGCTGTTGAGATGGCTGAGGAGAGAAACAGTCCCGGATAAG 1320
Db 1261 GGGAGAGCAGCAACAGGCGCTGTTGAGATGGCTGAGGAGAGAAACAGTCCCGGATAAG 1320
QY 1321 AACTGGAGGCGCTGCAGGTAAAAATCCACGGCTGGAGAGCTGTCGGGCGACTGCAGA 1380
Db 1321 AACTGGAGGCGCTGCAGGTAAAAATCCACGGCTGGAGAGCTGTCGGGCGACTGCAGA 1380
QY 1381 CAGAGCGCAATGACTGAAACAGAGGATACAGGACCTGAGTGTGTGGTGGCAGGCGTCCC 1440
Db 1381 CAGAGCGCAATGACTGAAACAGAGGATACAGGACCTGAGTGTGTGGTGGCAGGCGTCCC 1440
QY 1441 TCACTGACAGTGGCGCTTGAGAGAGGCGCAGAGGCGCTGGGCTCAAGACCCAGCTCCC 1500
Db 1441 TCACTGACAGTGGCGCTTGAGAGAGGCGCAGAGGCGCTGGGCTCAAGACCCAGCTCCC 1500
QY 1501 CCAGGCTCAGAGAGCGCTTGTACCCAGAGCAGCGAGCAGAGAGCATCAGGCGCAGA 1560
Db 1501 CCAGGCTCAGAGAGCGCTTGTACCCAGAGCAGCGAGCAGAGAGCATCAGGCGCAGA 1560

QY 1561 CTGGGCTTCAAGAGCCCACTCCGCGCAGGCGCTAGAGAGCCTGGTGTGGGTGATGCTGG 1620
Db 1561 CTGGGCTTCAAGAGCCCACTCCGCGCAGGCGCTAGAGAGCCTGGTGTGGGTGATGCTGG 1620
QY 1621 GAAGGAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1680
Db 1621 GAAGGAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1680
QY 1681 TTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTGGACATTCGCAATTTGGATTTT 1740
Db 1681 TTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTGGACATTCGCAATTTGGATTTT 1740
QY 1741 GTGGGTCACTTTTACGTACATAGGGCATTTTGAAGGCTTGCAGGCTTGCAGTATATACCTG 1800
Db 1741 GTGGGTCACTTTTACGTACATAGGGCATTTTGAAGGCTTGCAGGCTTGCAGTATATACCTG 1800
QY 1801 TAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
Db 1801 TAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
QY 1861 TCTGTAGCTGAAGAGCTTGTAGAGGGGCTGTCTATCTGTAGCTGCCATCAGTGAAGTTG 1920
Db 1861 TCTGTAGCTGAAGAGCTTGTAGAGGGGCTGTCTATCTGTAGCTGCCATCAGTGAAGTTG 1920
QY 1921 GCAGAGTGACTTGCAGCATTTCTCTGTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980
Db 1921 GCAGAGTGACTTGCAGCATTTCTCTGTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980
QY 1981 CAGAGCTCAAAACAAAGTAATACACCAAGGCTTTCACCTGCATTTCTTTGTGAGCAGGCT 2040
Db 1981 CAGAGCTCAAAACAAAGTAATACACCAAGGCTTTCACCTGCATTTCTTTGTGAGCAGGCT 2040
QY 2041 TGCTTGGTCACTCAGGCGCTCTAGCTGTGTGAGGCTCCTTTGATTTCTAGACCTGG 2100
Db 2041 TGCTTGGTCACTCAGGCGCTCTAGCTGTGTGAGGCTCCTTTGATTTCTAGACCTGG 2100
QY 2101 AAAAGGTCTCCCTAGCAGAGCCTTGCGAGGCGCTCAGAGCTGGGATTTCTCCGCTTGA 2160
Db 2101 AAAAGGTCTCCCTAGCAGAGCCTTGCGAGGCGCTCAGAGCTGGGATTTCTCCGCTTGA 2160
QY 2161 ACAAGGACCTTGGAGATGTTTTTGGGTGGGATGATGTGCTGGTTCAGAGGCGCTTGGG 2220
Db 2161 ACAAGGACCTTGGAGATGTTTTTGGGTGGGATGATGTGCTGGTTCAGAGGCGCTTGGG 2220
QY 2221 ATCGCTTCCCTGCCCTTGGTGTAGTCCAGGACAGGCGCAATGATGCTTCTCAGTAGCCT 2280
Db 2221 ATCGCTTCCCTGCCCTTGGTGTAGTCCAGGACAGGCGCAATGATGCTTCTCAGTAGCCT 2280
QY 2281 TATCAATCAGAGTGCCTCTTAGCCTGCACAAAATGATGACAGAGATCAACCAAGGA 2340
Db 2281 TATCAATCAGAGTGCCTCTTAGCCTGCACAAAATGATGACAGAGATCAACCAAGGA 2340
QY 2341 TTATTTCTGAGGCTGTTTTTTCTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2400
Db 2341 TTATTTCTGAGGCTGTTTTTTCTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2400
QY 2401 TTTTTTGCATGACAGTGTGTTTTGATTTGAGGACCTTCCAGGAAAAGGATGCTGTACC 2460
Db 2401 TTTTTTGCATGACAGTGTGTTTTGATTTGAGGACCTTCCAGGAAAAGGATGCTGTACC 2460
QY 2461 AGTGTGCTGGTGGCTTGGCTCCAGTGTCCAGCTCTTCCACCCAGCTTGGCTCC 2520
Db 2461 AGTGTGCTGGTGGCTTGGCTCCAGTGTCCAGCTCTTCCACCCAGCTTGGCTCC 2520
QY 2521 TTTTGCATCTTGATGCTCAGGTTTTCTGTTGGTGAGATCAGGTTGTTGTTGTTGTTGTT 2580
Db 2521 TTTTGCATCTTGATGCTCAGGTTTTCTGTTGGTGAGATCAGGTTGTTGTTGTTGTTGTT 2580
QY 2581 AAGGAAAAGGCTTCTGATGGCTTTGCCAAGCTTTACCTGTGGGTTTTCAGTCTCCTGAGG 2640
Db 2581 AAGGAAAAGGCTTCTGATGGCTTTGCCAAGCTTTACCTGTGGGTTTTCAGTCTCCTGAGG 2640
QY 2641 CCACCACAGTTCCCATCAGCACTGCTCTCCATGACGAGCTTCTGGGTCCCATGCTCCAGC 2700

Db 2641 CCACCACAGTTCCTCATCAGCACTGTCTCCATGACAGTGTCTGGGTGCCATGTCCAGC 2700
QY 2701 TGCCCTCTTGGCTTCATGGGTTTCTCTCTCTGCTCCGCCCAACCCACATGTGCAATCCT 2760
Db 2701 TGCCCTCTTGGCTTCATGGGTTTCTCTCTCTGCTCCGCCCAACCCACATGTGCAATCCT 2760
QY 2761 CAAGATTGTCTCTGATCTATTTCTCTGGCACCTCCCTGCTCTCTCTGGGGATTCCTACTT 2820
Db 2761 CAAGATTGTCTGATCTATTTCTCTGGCACCTCCCTGCTCTCTCTGGGGATTCCTACTT 2820
QY 2821 CTTCTCTGTGGGGCCCATAGCTTGTCTAAAGGTAAAGAAATGAATGAATCTATGTA 2880
Db 2821 CTTCTCTGTGGGGCCCATAGCTTGTCTAAAGGTAAAGAAATGAATGAATCTATGTA 2880
QY 2881 CTGGGGCCCGAGAAATCCATAGCTTGTCTAAAGGTAAAGAAATGAATGAATCTATGTA 2940
Db 2881 CTGGGGCCCGAGAAATCCATAGCTTGTCTAAAGGTAAAGAAATGAATGAATCTATGTA 2940
QY 2941 CCACTCCAGTACATACTATCTATCTATCTGTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
Db 2941 CCACTCCAGTACATACTATCTATCTATCTGTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
QY 3001 CTGTTGCAAACTTCGAGGGTATAGCCAGTGTGTGCGCAAGCAGTTTTCGGGGACAAC 3060
Db 3001 CTGTTGCAAACTTCGAGGGTATAGCCAGTGTGTGCGCAAGCAGTTTTCGGGGACAAC 3060
QY 3061 AGAATGACTCAGACCAAGATGGATAGATGGTGTAGGGCTTTGCTGTGCTGTTTCTT 3120
Db 3061 AGAATGACTCAGACCAAGATGGATAGATGGTGTAGGGCTTTGCTGTGCTGTTTCTT 3120
QY 3121 TGAATAGTCAATGCTCTGAGGTCCTTCACTTCCATACCTAGGCCACTCTTTTAGCC 3180
Db 3121 TGAATAGTCAATGCTCTGAGGTCCTTCACTTCCATACCTAGGCCACTCTTTTAGCC 3180
QY 3181 CTTACCTTAATCTCTCAGATAGTGTGTTTCAAGAAATGTTAAGTACTGAATCATGTG 3240
Db 3181 CTTACCTTAATCTCTCAGATAGTGTGTTTCAAGAAATGTTAAGTACTGAATCATGTG 3240
QY 3241 TGACTGAGACAGAGATGGCAATGAATGAATGCAACCAATTTCTCTCTCTGCTCCCGAGG 3300
Db 3241 TGACTGAGACAGAGATGGCAATGAATGAATGCAACCAATTTCTCTCTCTGCTCCCGAGG 3300
QY 3301 CAGGTACCACTGATCTGCATCAGAGTGTGCTGATTTCTCTGCTGATCTCTTCAATCTA 3360
Db 3301 CAGGTACCACTGATCTGCATCAGAGTGTGCTGATTTCTCTGCTGATCTCTTCAATCTA 3360
QY 3361 GGTGCCCTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTG 3420
Db 3361 GGTGCCCTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTG 3420
QY 3421 TGCTCTGTGAAGTGTTCCTCATGACCTTTTCTCTGCTTGAATCCCTTGTCTGAGAG 3480
Db 3421 TGCTCTGTGAAGTGTTCCTCATGACCTTTTCTCTGCTTGAATCCCTTGTCTGAGAG 3480
QY 3481 TAGTCTTGTCTCTCTCTGCTCCAGTAGGCTTTTCTTACCCAGGCTTTGTGCCAGG 3540
Db 3481 TAGTCTTGTCTCTCTCTGCTCCAGTAGGCTTTTCTTACCCAGGCTTTGTGCCAGG 3540
QY 3541 CTAAGCTGTGTAAGAGCTCCCACTCAAGAGTTTGTAGGCGAGAGAGTGTGAGGGA 3600
Db 3541 CTAAGCTGTGTAAGAGCTCCCACTCAAGAGTTTGTAGGCGAGAGAGTGTGAGGGA 3600
QY 3601 AGAGGCGAGAGTATGACCTTTCCCTTTGAAGAGAGGGGAAGCCCTACAGTGGCCCA 3660
Db 3601 AGAGGCGAGAGTATGACCTTTCCCTTTGAAGAGAGGGGAAGCCCTACAGTGGCCCA 3660
QY 3661 TAATTCCTGACTCACAATTCAGTACCTCTTAATGCTGTGAGGGAATGAGTGTGCTG 3720
Db 3661 TAATTCCTGACTCACAATTCAGTACCTCTTAATGCTGTGAGGGAATGAGTGTGCTG 3720
QY 3721 GATCCCAAGTGTGTGAGGAGGCAAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTG 3780

Db 3721 GATCCCAAGTGTGTGAGGAGGCAAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTG 3780
QY 3781 GGTCTAGGATGTGGGCCCCCAGGCAAGGTGACGCTTTGCTCAGACCTCCATCCATGCTA 3840
Db 3781 GGTCTAGGATGTGGGCCCCCAGGCAAGGTGACGCTTTGCTCAGACCTCCATCCATGCTA 3840
QY 3841 GACCTTTAGGGCCAGTCTGCAGATGAGTTCCTACTCTTTTCTCTCTTCTTCTTCTT 3900
Db 3841 GACCTTTAGGGCCAGTCTGCAGATGAGTTCCTACTCTTTTCTCTCTTCTTCTTCTT 3900
QY 3901 TCAACCAATCACTACAGTGTCTGCTTCTGCTTTTCAAGTAGCCAGGTCCTGGGCCA 3960
Db 3901 TCAACCAATCACTACAGTGTCTGCTTCTGCTTTTCAAGTAGCCAGGTCCTGGGCCA 3960
QY 3961 GATGAGGGAGGTGCTTATCCATGAGTGAAGGCGAGTGTCTTCTCACCCTGGGTGCTC 4020
Db 3961 GATGAGGGAGGTGCTTATCCATGAGTGAAGGCGAGTGTCTTCTCACCCTGGGTGCTC 4020
QY 4021 CACACTTGTGACCCCTCAGTTTGTAGACCCCAAGATCTGTGTTGTTTCTTGTAGTTGCTAGC 4080
Db 4021 CACACTTGTGACCCCTCAGTTTGTAGACCCCAAGATCTGTGTTGTTTCTTGTAGTTGCTAGC 4080
QY 4081 TTTTCTCCAGGGGACACACAGCAGTGAAGCTCAAGAGCGCATGCTCTGCTAATATGTA 4140
Db 4081 TTTTCTCCAGGGGACACACAGCAGTGAAGCTCAAGAGCGCATGCTCTGCTAATATGTA 4140
QY 4141 ATTGTTTTCAGGGCTTGTCCAGCTGAGAGCTTCAATGTCCACCAAGATTTCTGAGAGGTGTC 4200
Db 4141 ATTGTTTTCAGGGCTTGTCCAGCTGAGAGCTTCAATGTCCACCAAGATTTCTGAGAGGTGTC 4200
QY 4201 AGCAGCACTTTTCTTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 4260
Db 4201 AGCAGCACTTTTCTTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 4260
QY 4261 GAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTTATTTAAACCTCC 4320
Db 4261 GAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTTATTTAAACCTCC 4320
QY 4321 TCCACCCCATCAGCGTGGCCCTGAGGGCTGACCCGAGGCGCAGTGGAGCTGCTGTGTC 4380
Db 4321 TCCACCCCATCAGCGTGGCCCTGAGGGCTGACCCGAGGCGCAGTGGAGCTGCTGTGTC 4380
QY 4381 CACGGGGAGGCGCAAGGCTGCTGAGTGTATTTCCAGCTGTGCTCCCGAGCCTTTCCGC 4440
Db 4381 CACGGGGAGGCGCAAGGCTGCTGAGTGTATTTCCAGCTGTGCTCCCGAGCCTTTCCGC 4440
QY 4441 CTTGCACAGCACAGAGTGTGTCACCCAGGACAGCCAGGCACTGTCTCTTTGCTT 4500
Db 4441 CTTGCACAGCACAGAGTGTGTCACCCAGGACAGCCAGGCACTGTCTCTTTGCTT 4500
QY 4501 CCTGGGGAGAGGAGCTGCTTCTGCTCCCTGTAATCTTCTTATGCTTATGCTTATGCTT 4560
Db 4501 CCTGGGGAGAGGAGCTGCTTCTGCTCCCTGTAATCTTCTTATGCTTATGCTTATGCTT 4560
QY 4561 CACTCAGACTTTGTTGAGCTGCACTGGCAGCTTTTGTCTCTCTTTGGGTATTCACAC 4620
Db 4561 CACTCAGACTTTGTTGAGCTGCACTGGCAGCTTTTGTCTCTCTTTGGGTATTCACAC 4620
QY 4621 AGCCAGGAGCTGATTTTGTATGTTTAAACCAATTAATAAGAGTCTGTGCTT 4680
Db 4621 AGCCAGGAGCTGATTTTGTATGTTTAAACCAATTAATAAGAGTCTGTGCTT 4680
QY 4681 AAAAAAAAAAAAAA 4697
Db 4681 AAAAAAAAAAAAAA 4697

RESULT 3
AR409323
LOCUS
DEFINITION
ACCESSION
VERSION

AR409323
Sequence 17 from patent US 6632923.
AR409323
AR409323.1 GI:40160111
linear PAT 18-DEC-2003

Qy	1921	GCAGAAGTGACTTGAGCATTTCTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCTTTT	1981
Db	1921	GCAGAAGTGACTTGAGCATTTCTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCTTTT	1980
Qy	1981	CAGAGCTCAAAACAAGTAATAACCAACCAAGTCTTGACCTGCATTTGTCTTGTGACGAGGCT	2040
Db	1981	CAGAGCTCAAAACAAGTAATAACCAACCAAGTCTTGACCTGCATTTGTCTTGTGACGAGGCT	2040
Qy	2041	TGCTTGGTTCAGCTCAGGCGCTCTAGCTGCTTGGAGGCTCCTTTGATTTCTTAGACCTGG	2100
Db	2041	TGCTTGGTTCAGCTCAGGCGCTCTAGCTGCTTGGAGGCTCCTTTGATTTCTTAGACCTGG	2100
Qy	2101	AAAAGGTGTCCTTAGGCAGAGCCCTGCAGAGGCGCTCAGAGCTGGGATTTCTGCTGGGA	2160
Db	2101	AAAAGGTGTCCTTAGGCAGAGCCCTGCAGAGGCGCTCAGAGCTGGGATTTCTGCTGGGA	2160
Qy	2161	ACAAGGACCTGGAGAATGTTTTTGCCTGGGATGATGTCTGTGTCAGGAGCCCTTTGGGC	2220
Db	2161	ACAAGGACCTGGAGAATGTTTTTGCCTGGGATGATGTCTGTGTCAGGAGCCCTTTGGGC	2220
Qy	2221	ATCGCTTCCCTGCGCTTGTGTAGTGCAGGACCAAGCCCAATGATGCTTCTAGTAGCCT	2280
Db	2221	ATCGCTTCCCTGCGCTTGTGTAGTGCAGGACCAAGCCCAATGATGCTTCTAGTAGCCT	2280
Qy	2281	TATCATTTACAGAGTGCCCTCTAGACCTGCACAAATGATTCAGAGAGATCACCCAAAGGA	2340
Db	2281	TATCATTTACAGAGTGCCCTCTAGACCTGCACAAATGATTCAGAGAGATCACCCAAAGGA	2340
Qy	2341	TTATTTCTGAAGGTGTTTTTCTTTTAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2400
Db	2341	TTATTTCTGAAGGTGTTTTTCTTTTAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2400
Qy	2401	TTTTTTTTGCACATCAGAGTGTTTCTGATTAGGAGCCCTTCCAAGGAAAAGGATGCTGTACC	2460
Db	2401	TTTTTTTTGCACATCAGAGTGTTTCTGATTAGGAGCCCTTCCAAGGAAAAGGATGCTGTACC	2460
Qy	2461	AGTGGTGCTGGGTGCTGCGCTCCAGTGTCCACACTCTTTCACCACCCACATTTGCTGCC	2520
Db	2461	AGTGGTGCTGGGTGCTGCGCTCCAGTGTCCACACTCTTTCACCACCCACATTTGCTGCC	2520
Qy	2521	TTTTGCCATCTTGATGCTGAGGTTTTCTCTGTTTGGTGAGATCAGGTTGTTTGTGTAAGA	2580
Db	2521	TTTTGCCATCTTGATGCTGAGGTTTTCTCTGTTTGGTGAGATCAGGTTGTTTGTGTAAGA	2580
Qy	2581	AAGGAAAAGGGCTTCGTATGGCTTTGGCCACAAGCTTACTGTGGGTTTCAGTCCCTGAGAGG	2640
Db	2581	AAGGAAAAGGGCTTCGTATGGCTTTGGCCACAAGCTTACTGTGGGTTTCAGTCCCTGAGAGG	2640
Qy	2641	CCACCACAGTTCCCATCAGCAGTCTCCATGACAGAGTTGCTGGGTCCCATGTCCAGC	2700
Db	2641	CCACCACAGTTCCCATCAGCAGTCTCCATGACAGAGTTGCTGGGTCCCATGTCCAGC	2700
Qy	2701	TGCTCTTTGGCTTCATGGGTTTTTCTGCTTCCGCCCCCACCCCCACATGTGCATTCCT	2760
Db	2701	TGCTCTTTGGCTTCATGGGTTTTTCTGCTTCCGCCCCCACCCCCACATGTGCATTCCT	2760
Qy	2761	CAAGATTTGCTCGATTTCTATTTCTCGGACCTCCCTGCTGCTTGGGATTTCTACTTT	2820
Db	2761	CAAGATTTGCTCGATTTCTATTTCTCGGACCTCCCTGCTGCTTGGGATTTCTACTTT	2820
Qy	2821	CTTCTGTGGGGCCCATAGCTGTTGCTTAAACAGGTAAGAAATGAATTAATTAATGA	2880
Db	2821	CTTCTGTGGGGCCCATAGCTGTTGCTTAAACAGGTAAGAAATGAATTAATTAATGA	2880
Qy	2881	CTGGGCCCCAGAAATCCATAAAATGGCTGCAGACAGTTGTTTCTGTGTCCTGTTCTACCC	2940
Db	2881	CTGGGCCCCAGAAATCCATAAAATGGCTGCAGACAGTTGTTTCTGTGTCCTGTTCTACCC	2940
Qy	2941	CCACTCCAGTACATAACTATGTACTGTGTAGAGCCATTATATGCTGTAATGTTCTG	3000
Db	2941	CCACTCCAGTACATAACTATGTACTGTGTAGAGCCATTATATGCTGTAATGTTCTG	3000

3001	CTGTTGCAAACTTGTCCAGGGTAATTAGCCAGTGTTCGCAAGCAGTTCCTCGGGACAAC	3061	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT
3001	CTGTTGCAAACTTGTCCAGGGTAATTAGCCAGTGTTCGCAAGCAGTTCCTCGGGACAAC	3061	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT
3061	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT	3120	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT
3061	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT	3120	AGAAAGACTCAGACCAAGATGAGATGGTTTATGAGGCTTCCTCTCTTGGCTGTTTCTT
3121	TGAACCTAGTCAATTCCTGTCAGGTCCTCTTCACTCTCCATPACTAGCCCACTCTTTTAGCC	3180	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG
3121	TGAACCTAGTCAATTCCTGTCAGGTCCTCTTCACTCTCCATPACTAGCCCACTCTTTTAGCC	3180	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG
3181	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG	3240	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG
3181	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG	3240	CTTACTCTTAATCTCTCAGATAGTTCGTTTCACAAGAATGTTAAGTACTGAATCATGTG
3241	TGACTGAGACAGAGATGGCAAAATGAATGGCACACCAATTCCTCTCTCTGTCGCCAGGG	3300	TGACTGAGACAGAGATGGCAAAATGAATGGCACACCAATTCCTCTCTCTGTCGCCAGGG
3241	TGACTGAGACAGAGATGGCAAAATGAATGGCACACCAATTCCTCTCTCTGTCGCCAGGG	3300	TGACTGAGACAGAGATGGCAAAATGAATGGCACACCAATTCCTCTCTCTGTCGCCAGGG
3301	CAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATCTCTCGGTGATCCTTCACTCTA	3360	CAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATCTCTCGGTGATCCTTCACTCTA
3301	CAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATCTCTCGGTGATCCTTCACTCTA	3360	CAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATCTCTCGGTGATCCTTCACTCTA
3361	GGTGCCCTCAAGCAGCTGTGTGAGTGTCTGCGCATCTCTGCGCTCAGATACTGCG	3420	GGTGCCCTCAAGCAGCTGTGTGAGTGTCTGCGCATCTCTGCGCTCAGATACTGCG
3361	GGTGCCCTCAAGCAGCTGTGTGAGTGTCTGCGCATCTCTGCGCTCAGATACTGCG	3420	GGTGCCCTCAAGCAGCTGTGTGAGTGTCTGCGCATCTCTGCGCTCAGATACTGCG
3421	TGCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCTGTGCTGGAG	3480	TGCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCTGTGCTGGAG
3421	TGCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCTGTGCTGGAG	3480	TGCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCTGTGCTGGAG
3481	TAGTCCTTGCCTCTCTTCTGCTCCAGTAGGCTTTTCTTACCAGCCCTTGTGCCAGG	3540	TAGTCCTTGCCTCTCTTCTGCTCCAGTAGGCTTTTCTTACCAGCCCTTGTGCCAGG
3481	TAGTCCTTGCCTCTCTTCTGCTCCAGTAGGCTTTTCTTACCAGCCCTTGTGCCAGG	3540	TAGTCCTTGCCTCTCTTCTGCTCCAGTAGGCTTTTCTTACCAGCCCTTGTGCCAGG
3541	CTAAGCTGGTACAAGAGCTGCCAACTCAGAGTTCCTGCTAGGCGAGAGGTCAGGGA	3600	CTAAGCTGGTACAAGAGCTGCCAACTCAGAGTTCCTGCTAGGCGAGAGGTCAGGGA
3541	CTAAGCTGGTACAAGAGCTGCCAACTCAGAGTTCCTGCTAGGCGAGAGGTCAGGGA	3600	CTAAGCTGGTACAAGAGCTGCCAACTCAGAGTTCCTGCTAGGCGAGAGGTCAGGGA
3601	AGAGCAGAGGTATGCACCTTCCCTTTGAAGAGGGGGAAGGCTACAGTGGCCACA	3660	AGAGCAGAGGTATGCACCTTCCCTTTGAAGAGGGGGAAGGCTACAGTGGCCACA
3601	AGAGCAGAGGTATGCACCTTCCCTTTGAAGAGGGGGAAGGCTACAGTGGCCACA	3660	AGAGCAGAGGTATGCACCTTCCCTTTGAAGAGGGGGAAGGCTACAGTGGCCACA
3661	TAAATGCCCTGACTCACAACCTTCACTACCTCTTAAATGCCCTGTGGAGGACTGAGCTGCTG	3720	TAAATGCCCTGACTCACAACCTTCACTACCTCTTAAATGCCCTGTGGAGGACTGAGCTGCTG
3661	TAAATGCCCTGACTCACAACCTTCACTACCTCTTAAATGCCCTGTGGAGGACTGAGCTGCTG	3720	TAAATGCCCTGACTCACAACCTTCACTACCTCTTAAATGCCCTGTGGAGGACTGAGCTGCTG
3721	GATCCAGTGTGTGTGTAAGAGGCCACAGTAGCAGAGTGGCCCCAGCTGGGTTCCCA	3780	GATCCAGTGTGTGTGTAAGAGGCCACAGTAGCAGAGTGGCCCCAGCTGGGTTCCCA
3721	GATCCAGTGTGTGTGTAAGAGGCCACAGTAGCAGAGTGGCCCCAGCTGGGTTCCCA	3780	GATCCAGTGTGTGTGTAAGAGGCCACAGTAGCAGAGTGGCCCCAGCTGGGTTCCCA
3781	GGTCAGGAATGTGGGCCCCAGGCAAGGTGCAGCCTTTGCTCACAGCTCCATCCATGTCCTA	3840	GGTCAGGAATGTGGGCCCCAGGCAAGGTGCAGCCTTTGCTCACAGCTCCATCCATGTCCTA
3781	GGTCAGGAATGTGGGCCCCAGGCAAGGTGCAGCCTTTGCTCACAGCTCCATCCATGTCCTA	3840	GGTCAGGAATGTGGGCCCCAGGCAAGGTGCAGCCTTTGCTCACAGCTCCATCCATGTCCTA
3841	GACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTACTTTTCTCTTCAATGACCAAA	3900	GACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTACTTTTCTCTTCAATGACCAAA
3841	GACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTACTTTTCTCTTCAATGACCAAA	3900	GACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTACTTTTCTCTTCAATGACCAAA
3901	TCAACCAATCACTACAGCTCTGTGTTTCTGTTTCCAAAGTAGCCAGGTCCTGGGCCA	3960	TCAACCAATCACTACAGCTCTGTGTTTCTGTTTCCAAAGTAGCCAGGTCCTGGGCCA
3901	TCAACCAATCACTACAGCTCTGTGTTTCTGTTTCCAAAGTAGCCAGGTCCTGGGCCA	3960	TCAACCAATCACTACAGCTCTGTGTTTCTGTTTCCAAAGTAGCCAGGTCCTGGGCCA
3961	GATGAGGGGAGGTGCCCTATCCATGAGTGAAGGCCAGTGTCTTCTCACTGGGTGGTCC	4020	GATGAGGGGAGGTGCCCTATCCATGAGTGAAGGCCAGTGTCTTCTCACTGGGTGGTCC
3961	GATGAGGGGAGGTGCCCTATCCATGAGTGAAGGCCAGTGTCTTCTCACTGGGTGGTCC	4020	GATGAGGGGAGGTGCCCTATCCATGAGTGAAGGCCAGTGTCTTCTCACTGGGTGGTCC
4021	CACACTTGTGACCTCAGTTTTTAGCCCAAGATCTGTGTTTCTTAGATTGCTAGC	4080	CACACTTGTGACCTCAGTTTTTAGCCCAAGATCTGTGTTTCTTAGATTGCTAGC
4021	CACACTTGTGACCTCAGTTTTTAGCCCAAGATCTGTGTTTCTTAGATTGCTAGC	4080	CACACTTGTGACCTCAGTTTTTAGCCCAAGATCTGTGTTTCTTAGATTGCTAGC
4081	TTTTTCTCCAGGGGACACAGCAGGTGAAGCTCAAGAGCGCATGGCTCTGCTAATAGTAA	4140	TTTTTCTCCAGGGGACACAGCAGGTGAAGCTCAAGAGCGCATGGCTCTGCTAATAGTAA

Db 4081 TTTTCTCCAGGGGACCAAGAGGTTGAAGCTCAAGAGGCGATGGCTCTGCTAAATAGTAA 4140
QY 4141 ATTGTTTTAGGGGCTTGTCCAGCTGAGAGCTTCATGTCACAGATTTCTGAGAGGTGTC 4200
Db 4141 ATTGTTTTAGGGGCTTGTCCAGCTGAGAGCTTCATGTCACAGATTTCTGAGAGGTGTC 4200
QY 4201 AGCAGACATTTTTTTTTTTTATTTGTTGTTGTTTCCATGAGGTTATCCGACCATGGGCT 4260
Db 4201 AGCAGACATTTTTTTTTTTTATTTGTTGTTGTTTCCATGAGGTTATCCGACCATGGGCT 4260
QY 4261 GAGCTCAGGACATTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTTAAACCTTCC 4320
Db 4261 GAGCTCAGGACATTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTTAAACCTTCC 4320
QY 4321 TCCACCCCATCAGGTTGGCCCTCAGGCTGACCCGAGGCGCAGTGAGCTGCTGTGTGTC 4380
Db 4321 TCCACCCCATCAGGTTGGCCCTCAGGCTGACCCGAGGCGCAGTGAGCTGCTGTGTGTC 4380
QY 4381 CACGGGGAGGGGCAAGGCTGCTGAGCTGATTTCTCCAGCTGTGCCCCAGGCTTTCCGC 4440
Db 4381 CACGGGGAGGGGCAAGGCTGCTGAGCTGATTTCTCCAGCTGTGCCCCAGGCTTTCCGC 4440
QY 4441 CTTGCAAGACAGAGGTTGTACCCCGAGGACGCGACCTGCTCTTGCCTT 4500
Db 4441 CTTGCAAGACAGAGGTTGTACCCCGAGGACGCGACCTGCTCTTGCCTT 4500
QY 4501 CTTGCGGGAAGGAGCTGCTTCTGCTTCTGTAACCTGTTTCTTATGCCCCAACCCCGC 4560
Db 4501 CTTGCGGGAAGGAGCTGCTTCTGCTTCTGTAACCTGTTTCTTATGCCCCAACCCCGC 4560
QY 4561 CACTCAGACTTTTGAAGCTGACCTGCGAGCTTTTGTCTCTTGGTATTCACAAC 4620
Db 4561 CACTCAGACTTTTGAAGCTGACCTGCGAGCTTTTGTCTCTTGGTATTCACAAC 4620
QY 4621 AGCCAGGACTGATTTGATGATTTTAAACACATTAATAAGAGTCTGTGCTTA 4680
Db 4621 AGCCAGGACTGATTTGATGATTTTAAACACATTAATAAGAGTCTGTGCTTA 4680
QY 4681 AAAAAAAAAAAAAA 4697
Db 4681 AAAAAAAAAAAAAA 4697

RESULT 4
AX239574
LOCUS
DEFINITION
Sequence 17 from Patent WO0164874.
ACCESSION
AX239574
VERSION
AX239574.1
KEYWORDS
GI:15797259
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Lees A.M., Lees R.S., Law S.W. and Arjona A.A.
Low density lipoprotein binding proteins and their use in diagnosing
and treating atherosclerosis
Patent: WO 0164874-A 17 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
JOURNAL
Location/Qualifiers
FEATURES
source
1..4697
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1..1595
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC88297.1"
/db_xref="GI:15797260"
/db_xref="REMBL:CA88297"
/translation="KSSFGQPEAGPEGAQERPSQAPVAEAGFGSSQAPRKPEGQAQ"

RTAOSGALRDVSEELSLEQLIEDILSTYCVNNDNNGGPGEDGAGBPAPPEDEAKSRTYVA
RNEPTPTPVYGEKEPSKGDPTNEIRQDEVDHRRPQKKAQKAGLGGKILLM
QTLNLTSTPEKLAALCKYAELEHNSQKQMLIKQSQSLQVQEKHLRGEHSA
VLARKLESRLORHNRSLKEBVGORAREEERKEVTSHFQVTLNDLQLOMECHN
FRSKLQENMELAEARLKLIEOVELSEBHIKDFVFKDLQOOLVDAKLOAOEMLKE
AERHQBKDFLLKEAVESORMELMKOQTHLKOOLALYTEKPEFNTLSQKSEVFE
TTFQEMKMTKKIKLEKETITWRSWESSNKALLEMAEBEKTVRDELSGLQVKIOR
LEKLCRALQTERNDLNKRVQDLGAGGGGSLTDSGPERPPEGPGAFSPFRVTEAPCI
PGAPSTBASGQTGPQETTSARA"
ORIGIN
Query Match 100.0%; Score 4697; DB 6; Length 4697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAAAAGCAGCCAGGCAACCCGGAAGCAGGACCCGAGGGAGCCAGAGGCGGCCAGCC 60
Db 1 CAAAAAGCAGCCAGGCAACCCGGAAGCAGGACCCGAGGGAGCCAGAGGCGGCCAGCC 60
QY 61 AGGCGGCTCTCGAGTAGAAGCAGAGGTCCTCCGCGACGACGCGCTCTCGGAAGCGG 120
Db 61 AGGCGGCTCTCGAGTAGAAGCAGAGGTCCTCCGCGACGACGCGCTCTCGGAAGCGG 120
QY 121 AGGGGCTCAAGCCAGAACGCTCAGTCTGGGGGCCCTTCGTGATGTCTCTGAGGAGTGA 180
Db 121 AGGGGCTCAAGCCAGAACGCTCAGTCTGGGGGCCCTTCGTGATGTCTCTGAGGAGTGA 180
QY 181 GCGGCAACTGGAAGACATCTGAGCAGATCTGTGGACAATAACAGAGGGGGGCCCG 240
Db 181 GCGGCAACTGGAAGACATCTGAGCAGATCTGTGGACAATAACAGAGGGGGGCCCG 240
QY 241 GCGGAGTGGGCAAGGCTGAGCGGCTGAACCCGAGATGTCAGAGAGTCCCGGAGCT 300
Db 241 GCGGAGTGGGCAAGGCTGAGCGGCTGAACCCGAGATGTCAGAGAGTCCCGGAGCT 300
QY 301 ATGTGGCAAGGATGGGAGCTGAAACCACTCCAGTATCTATGAGAGAGGAGAACCT 360
Db 301 ATGTGGCAAGGATGGGAGCTGAAACCACTCCAGTATCTATGAGAGAGGAGAACCT 360
QY 361 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGTGACGAGTCCGAGACCGAGACC 420
Db 361 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGTGACGAGTCCGAGACCGAGACC 420
QY 421 ATCGAAGGCCACAG 480
Db 421 ATCGAAGGCCACAG 480
QY 481 TGCAGACATTTGAATCTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TGCAGACATTTGAATCTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 ATGCTGAACTGCTGGAGGAGACCCGGAATTCACAGAAGCAGATGAAGTCTCTACAGAAA 600
Db 541 ATGCTGAACTGCTGGAGGAGACCCGGAATTCACAGAAGCAGATGAAGTCTCTACAGAAA 600
QY 601 AGCAGAGCAGCTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 AGCAGAGCAGCTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 TGGCCCGCAGCAGCTGTGAGGCTATGCGGTGAGTCCAGGCGACCAACCTCCCTCA 720
Db 661 TGGCCCGCAGCAGCTGTGAGGCTATGCGGTGAGTCCAGGCGACCAACCTCCCTCA 720
QY 721 AGGAAGAGAGCTGTGAGGCGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 AGGAAGAGAGCTGTGAGGCGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 ACTTCAGGTGACATGATGATTCAGTCCAGATGAGACACACATGAGCCCACT 840
Db 781 ACTTCAGGTGACATGATGATTCAGTCCAGATGAGACACACATGAGCCCACT 840
QY 841 CCAAGTGGCGCAAGAGAGACATGGAGCTGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
Db 841 CCAAGTGGCGCAAGAGAGACATGGAGCTGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900

Db 841 CCAAGCTGGCCAAAGAGAAATCATGAGCTGGCTGAGAGGCTCAAGAACTGATTGAGCAGT 900
Qy 901 ATGAGCTGGCGAGAGAGATATCCAAAGTCTTCAACACAGAGAGCTTACACAGCAGC 960
Db 901 ATGAGCTGGCGAGAGAGATATCCAAAGTCTTCAACACAGAGAGCTTACACAGCAGC 960
Qy 961 TGGTGGATGCCAAGCTCCAGCAGAGGCCAGGAGATGCTAAAGGAGGCAAGAGAGCGGCACC 1020
Db 961 TGGTGGATGCCAAGCTCCAGCAGAGGCCAGGAGATGCTAAAGGAGGCAAGAGAGCGGCACC 1020
Qy 1021 AGCGGAGAGAGATTTCTCTGAAAGAGCGAGTAGAGTCCAGAGAGATGTGTAGCTGA 1080
Db 1021 AGCGGAGAGAGATTTCTCTGAAAGAGCGAGTAGAGTCCAGAGAGATGTGTAGCTGA 1080
Qy 1081 TGAAGCAGCAGAGAGACCCACCTGAAGCAACAGCTTGGCCCTATACACAGAGAGATTTGAGG 1140
Db 1081 TGAAGCAGCAGAGAGACCCACCTGAAGCAACAGCTTGGCCCTATACACAGAGAGATTTGAGG 1140
Qy 1141 AGTTCCAGAACACACTTTCCAAAAGCAGCGAGGTATTCCACCAATTCAAGCAGAGATGG 1200
Db 1141 AGTTCCAGAACACACTTTCCAAAAGCAGCGAGGTATTCCACCAATTCAAGCAGAGATGG 1200
Qy 1201 AAGAGATGACTAAGAGATCAAGAGCTGGAGAGAGAAACCAACATGTACCGCTCCCGT 1260
Db 1201 AAGAGATGACTAAGAGATCAAGAGCTGGAGAGAGAAACCAACATGTACCGCTCCCGT 1260
Qy 1261 GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAGAAACCAAGTCCGGGATAAAG 1320
Db 1261 GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAGAAACCAAGTCCGGGATAAAG 1320
Qy 1321 AACTGGAGGGCTCGAGGTAAATCCAAAGCTGGAGAGAGCTGTGCGGGGACTGGAGA 1380
Db 1321 AACTGGAGGGCTCGAGGTAAATCCAAAGCTGGAGAGAGCTGTGCGGGGACTGGAGA 1380
Qy 1381 CAGAGCGCAATGACTGAAAGAGGGTACAGGACCTGAGTGTGGTGGCCAGCGCTCC 1440
Db 1381 CAGAGCGCAATGACTGAAAGAGGGTACAGGACCTGAGTGTGGTGGCCAGCGCTCC 1440
Qy 1441 TCACGTGACGTGGCCCTGAGAGAGGGCCAGAGGGCCCTGGGGCTCAAGCACCCAGCTCC 1500
Db 1441 TCACGTGACGTGGCCCTGAGAGAGGGCCAGAGGGCCCTGGGGCTCAAGCACCCAGCTCC 1500
Qy 1501 CCAGGCTCAGAGAGCGCTTGTACCCAGGAGCAGCAGCAGCAGAGCATCAGGCGAGA 1560
Db 1501 CCAGGCTCAGAGAGCGCTTGTACCCAGGAGCAGCAGCAGCAGAGCATCAGGCGAGA 1560
Qy 1561 CTGGCCCTCAAGAGCCCACTCCGCCAGGCGCTAGAGAGCCTGGTGTGGGTCAATGCTGG 1620
Db 1561 CTGGCCCTCAAGAGCCCACTCCGCCAGGCGCTAGAGAGCCTGGTGTGGGTCAATGCTGG 1620
Qy 1621 GAGGGAGCGGAGCCAGCAGCGCTGGCCCATAAAGGCTCCCATGTGAGCAGCCCA 1680
Db 1621 GAGGGAGCGGAGCCAGCAGCGCTGGCCCATAAAGGCTCCCATGTGAGCAGCCCA 1680
Qy 1681 TTGCTGAAGCCAGGATGTTCTGACCTGGCTGGCATCTGGCACTTGCATTTTGGATTTT 1740
Db 1681 TTGCTGAAGCCAGGATGTTCTGACCTGGCTGGCATCTGGCACTTGCATTTTGGATTTT 1740
Qy 1741 GTGGGTGAGTACGTACATAGGCGATTTTGAAGCCCTTGAAGTGAATTTATACCTG 1800
Db 1741 GTGGGTGAGTACGTACATAGGCGATTTTGAAGCCCTTGAAGTGAATTTATACCTG 1800
Qy 1801 TAAGTGACAGTGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
Db 1801 TAAGTGACAGTGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
Qy 1861 TCTGTGAGCTGAAGAGTCTTTGAGAGGGCTGTGATCTGTAGCTGCCATCAGAGTGAAGTTG 1920
Db 1861 TCTGTGAGCTGAAGAGTCTTTGAGAGGGCTGTGATCTGTAGCTGCCATCAGAGTGAAGTTG 1920
Qy 1921 GCAGAGTGTACTGAGCATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980
Db 1921 GCAGAGTGTACTGAGCATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980

Qy 1981 CAGAGCTCAAAACAAGTAATACACAAAGTCTTGACTGATTTGCTTGTGAGCAGGCT 2040
Db 1981 CAGAGCTCAAAACAAGTAATACACAAAGTCTTGACTGATTTGCTTGTGAGCAGGCT 2040
Qy 2041 TGCCTTGGTCAAGCTCAGGCGCTCTTAGCTGCTTGGAGGCTCCTTTGATTTCTTAGACCTGG 2100
Db 2041 TGCCTTGGTCAAGCTCAGGCGCTCTTAGCTGCTTGGAGGCTCCTTTGATTTCTTAGACCTGG 2100
Qy 2101 AAAAGTGTCCCTAGCAGAGCCCTGGCAGGCGCTCAGAGCTGGAGATTTCTGCTGGA 2160
Db 2101 AAAAGTGTCCCTAGCAGAGCCCTGGCAGGCGCTCAGAGCTGGAGATTTCTGCTGGA 2160
Qy 2161 ACAAGGACCTCGAGAAATGTTTTGCGTGGAGATGCTGCTGGTCAAGAGCCCTTGGGC 2220
Db 2161 ACAAGGACCTCGAGAAATGTTTTGCGTGGAGATGCTGCTGGTCAAGAGCCCTTGGGC 2220
Qy 2221 ATCGCTTCCCTGCGCTTTGGTAGTGCAGGACCAAGGCAATGATGCTTCTCAGTAGCCT 2280
Db 2221 ATCGCTTCCCTGCGCTTTGGTAGTGCAGGACCAAGGCAATGATGCTTCTCAGTAGCCT 2280
Qy 2281 TATCATTACAGGTGCTCTCTAGCTGCACAAATGATTGACAGAGATCACCAGAGGA 2340
Db 2281 TATCATTACAGGTGCTCTCTAGCTGCACAAATGATTGACAGAGATCACCAGAGGA 2340
Qy 2341 TTATTTCTCAAGGTGTTTTTTCTTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTT 2400
Db 2341 TTATTTCTCAAGGTGTTTTTTCTTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTT 2400
Qy 2401 TTTTTTGCACATGACAGTGTGTTGATTTAGAGACCTTCCAAAGGAAAGGAGATGCTGACC 2460
Db 2401 TTTTTTGCACATGACAGTGTGTTGATTTAGAGACCTTCCAAAGGAAAGGAGATGCTGACC 2460
Qy 2461 AGTGTGCTCGGCTCGGCTCCAGTGTCCACCTCTTTCACACACCCACTTGGCTCC 2520
Db 2461 AGTGTGCTCGGCTCGGCTCCAGTGTCCACCTCTTTCACACACCCACTTGGCTCC 2520
Qy 2521 TTTGGCATCTTGTAGCTGAGGTTTCTGTTTGGTGGAGATCAGGTTGTTTGGTAAAGA 2580
Db 2521 TTTGGCATCTTGTAGCTGAGGTTTCTGTTTGGTGGAGATCAGGTTGTTTGGTAAAGA 2580
Qy 2581 AAGGAAAGGCTTCTGATGGCTTTGCCAACAGCTTACCTGGGTGTTTCAGTCTGAGAGG 2640
Db 2581 AAGGAAAGGCTTCTGATGGCTTTGCCAACAGCTTACCTGGGTGTTTCAGTCTGAGAGG 2640
Qy 2641 CCACCAACAGTCCCATCAGCAGTGTCTCCATGAGCAGATGCTGGGTCCCATGTCAGC 2700
Db 2641 CCACCAACAGTCCCATCAGCAGTGTCTCCATGAGCAGATGCTGGGTCCCATGTCAGC 2700
Qy 2701 TGCCTCTTTGGCTCATGGGTTTTTCTGCTTCCCTGCGCTGCTCTTGGGATTTCTACTT 2760
Db 2701 TGCCTCTTTGGCTCATGGGTTTTTCTGCTTCCCTGCGCTGCTCTTGGGATTTCTACTT 2760
Qy 2761 CAAGATTTGCTCCTGATTTCTATTTCTGCGACCTCCCTGCTGCTCTTGGGATTTCTACTT 2820
Db 2761 CAAGATTTGCTCCTGATTTCTATTTCTGCGACCTCCCTGCGCTGCTCTTGGGATTTCTACTT 2820
Qy 2821 CTTCTGTGTGGGCCCCATAGCTGTTCTTAAAGTAAAGAAATGAATTTGAATTTGA 2880
Db 2821 CTTCTGTGTGGGCCCCATAGCTGTTCTTAAAGTAAAGAAATGAATTTGAATTTGA 2880
Qy 2881 CTGGGCCCCAGAAATCCATAAATGGCTGCGACAGTGTGTTTCTGCTGCTGTTTCTACCC 2940
Db 2881 CTGGGCCCCAGAAATCCATAAATGGCTGCGACAGTGTGTTTCTGCTGCTGTTTCTACCC 2940
Qy 2941 CCACTCAGTACATACTATGATGCTAGAGCATTCTATATGCTGAATTTCTG 3000
Db 2941 CCACTCAGTACATACTATGATGCTAGAGCATTCTATATGCTGAATTTCTG 3000
Qy 3001 CTGTGCAAACTTCCAGGGTATTAGCAGTGTGTTGTCAGAGCAGTGTTCGGGCAAC 3060
Db 3001 CTGTGCAAACTTCCAGGGTATTAGCAGTGTGTTGTCAGAGCAGTGTTCGGGCAAC 3060

3061 AGAATGACTCAGACCAAGATGATAGGATGGTTAGGGCTTTGCTTCTGCTGTTTCTT 3120
3061 AGAATGACTCAGACCAAGATGATAGGATGGTTAGGGCTTTGCTTCTGCTGTTTCTT 3120
3121 TGAACCTAGTCATCTGCTGAGTGCCTTTCATCTCCATACCTAGGCCACTCTTTAGCC 3180
3121 TGAACCTAGTCATCTGCTGAGTGCCTTTCATCTCCATACCTAGGCCACTCTTTAGCC 3180
3181 CTTTACCTTAAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
3181 CTTTACCTTAAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
3241 TGAATGAGACCAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
3241 TGAATGAGACCAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
3301 CAGGTACCACTGATCTGCATCAGAGTTGCTGCTATTCTCTGCTGATCTCTTCAATCTA 3360
3301 CAGGTACCACTGATCTGCATCAGAGTTGCTGCTATTCTCTGCTGATCTCTTCAATCTA 3360
3361 GGTGCTCTCAGCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
3361 GGTGCTCTCAGCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
3421 TGTCTCTGTAAGTGTTCCTCATGACCTTTTCTTCCCTTTTGAATCCCTCTTGTCTGGAG 3480
3421 TGTCTCTGTAAGTGTTCCTCATGACCTTTTCTTCCCTTTTGAATCCCTCTTGTCTGGAG 3480
3481 TAGTCTCTGCTTCTTCTGCTCAGTAGGCTTCTTCTTCCCTTTTGAATCCCTCTTGTCTGGAG 3540
3481 TAGTCTCTGCTTCTTCTGCTCAGTAGGCTTCTTCTTCCCTTTTGAATCCCTCTTGTCTGGAG 3540
3541 CTAAGCTGTGACAGAGCTGCCAATCAGAGTTTGTGAGGAGGAGAGAGAGAGAGAGAGAG 3600
3541 CTAAGCTGTGACAGAGCTGCCAATCAGAGTTTGTGAGGAGGAGAGAGAGAGAGAGAGAG 3600
3601 AGAGGAGAGAGATGACCTTCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
3601 AGAGGAGAGAGATGACCTTCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
3661 TAATGCTGCTGACTCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 3720
3661 TAATGCTGCTGACTCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 3720
3721 GATCCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3780
3721 GATCCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3780
3781 GGTGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3840
3781 GGTGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3840
3841 GACCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3900
3841 GACCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3900
3901 TCAACCAATCAGTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960
3901 TCAACCAATCAGTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960
3961 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
3961 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
4021 CACACTTGTGACCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
4021 CACACTTGTGACCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
4081 TTTTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4140
4081 TTTTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4140
4141 ATTGTTTTCAGGGGCTTGTCTCAGCTGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 4200

4141 ATTGTTTTCAGGGGCTTGTCTCAGCTGAGAGCTTCACTCCACAGATCTCTGAGAGGTGC 4200
4201 AGCAGCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4260
4201 AGCAGCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4260
4261 GAGCTCAGGCACTTTCTGCTAGGAGCTGTTATTTCTGTAAGATGTTTATTTTAAACCTCC 4320
4261 GAGCTCAGGCACTTTCTGCTAGGAGCTGTTATTTCTGTAAGATGTTTATTTTAAACCTCC 4320
4321 TCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGG 4380
4321 TCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGG 4380
4381 CACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440
4381 CACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440
4441 CTTGACACAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4500
4441 CTTGACACAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4500
4501 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4560
4501 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4560
4561 CACTCAGAGCTTGTGTTGAAGCTGCACTGCGAGCTTTTGTCTCTTGTGGGTATTCACAA 4620
4561 CACTCAGAGCTTGTGTTGAAGCTGCACTGCGAGCTTTTGTCTCTTGTGGGTATTCACAA 4620
4621 AGCCAGGAGCTTGTATTTGATGATTTTAAACCACTTAAATAAAGAGCTGTGTTGCCCTTA 4680
4621 AGCCAGGAGCTTGTATTTGATGATTTTAAACCACTTAAATAAAGAGCTGTGTTGCCCTTA 4680
4681 AAAAAAATAAAAAA 4697
4681 AAAAAAATAAAAAA 4697

RESULT 5

BD056450

LOCUS

DEFINITION

Novel low density lipoprotein binding proteins and their use in

diagnosing and treating atherosclerosis.

ACCESSION

BD056450

VERSION

JP 2001506983-A/8

KEYWORDS

Aequorea victoria

ORGANISM

Aequorea victoria

REFERENCE

1 (bases 1 to 4697)

AUTHORS

Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.

TITLE

Novel low density lipoprotein binding proteins and their use in

diagnosing and treating atherosclerosis

JOURNAL

Patent: JP 2001506983-A 8 29-MAY-2001;

COMMENT

BOSTON HEART FOUNDATION INC

PN JP 2001506983-A/8

PF 29-MAY-2001

PR 26-NOV-1997 JP 1998524870

PI 60/031930, 03-JUN-1997 US 60/048547

ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC

A61K38/04, A61K38/17, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC

C07H21/00,

PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC

Strandedness: Single;

CC Topology: Linear;

EH Key Location/Qualifiers.

I. .4697

/organism="Aequorea victoria"

/mol_type="genomic DNA"

ORIGIN		/db_xref="taxon:6100"	
Query Match		100.0%;	Score 4697; DB 6; Length 4697;
Best Local Similarity		100.0%;	Pred. No. 0;
Matches 4697; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	CAAAAGCAGCCCGAGCAACCGGAGCAGACCGGAGGAGCCCGAGAGGCGCCAGCC	60
DB.	1	CAAAAGCAGCCCGAGCAACCGGAGCAGACCGGAGGAGCCCGAGAGGCGCCAGCC	60
QY	61	AGGCGCTCTCCAGTGAAGCAGAAAGGTCCCGCAGCAGCCAGGCTCTCTCGAAGCCGG	120
DB	61	AGGCGCTCTCCAGTGAAGCAGAAAGGTCCCGCAGCAGCCAGGCTCTCTCGAAGCCGG	120
QY	121	AGGGGCTCAAGCCAGAAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA	180
DB	121	AGGGGCTCAAGCCAGAAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA	180
QY	181	CCCCCRACTGGAGACATCTGAGCACAATCTGTGGGCAATAAACCAGGGGGCCCG	240
DB	181	CCCCCRACTGGAGACATCTGAGCACAATCTGTGGGCAATAAACCAGGGGGCCCG	240
QY	241	CGAGGATGGGCGACAGGGTGAAGCCGAGTGAACCCGAAGATGCAAGAACTCCCGGACCT	300
DB	241	CGAGGATGGGCGACAGGGTGAAGCCGAGTGAACCCGAAGATGCAAGAACTCCCGGACCT	300
QY	301	ATGTGGCAGGAATGGGAGGCTGAACCAACTCCAGTGTCTATGGAGAGAAAGACCT	360
DB	301	ATGTGGCAGGAATGGGAGGCTGAACCAACTCCAGTGTCTATGGAGAGAAAGACCT	360
QY	361	CAAAGGGGATCCAAACACAGAGATCCGGCAGAGTGAAGGTGGGGAAGGATCACGTGCTGA	420
DB	361	CAAAGGGGATCCAAACACAGAGATCCGGCAGAGTGAAGGTGGGGAAGGATCACGTGCTGA	420
QY	421	ATCGAAGCCACAGAGAGAAAGCAAGGGTTGGGGAAGGAGATCACGTGCTGA	480
DB	421	ATCGAAGCCACAGAGAGAAAGCAAGGGTTGGGGAAGGAGATCACGTGCTGA	480
QY	481	TGCAGACATTAATCTCTGATGATCCCGAGAGAGTGGTGTCTGTGCAAGAGT	540
DB	481	TGCAGACATTAATCTCTGATGATCCCGAGAGAGTGGTGTCTGTGCAAGAGT	540
QY	541	ATGCTGAATCTCTCGAGAGCACCGGAATTCACAGAGCAGATGAAGCTCCTACAGAAA	600
DB	541	ATGCTGAATCTCTCGAGAGCACCGGAATTCACAGAGCAGATGAAGCTCCTACAGAAA	600
QY	601	AGCAGCCAGCTGGTGCAGAGAGACCACTCCCGGTGAGCAGAGGCGCTCC	660
DB	601	AGCAGCCAGCTGGTGCAGAGAGACCACTCCCGGTGAGCAGAGGCGCTCC	660
QY	661	TGGCCCGCAGCAGCTTGAAGCCTATGCGGTGAGCTGCAGCGCACCAACCGCTCCCTCA	720
DB	661	TGGCCCGCAGCAGCTTGAAGCCTATGCGGTGAGCTGCAGCGCACCAACCGCTCCCTCA	720
QY	721	AGGAAGAGTGTGACGGGCGCGGAGAGAGAGAGCGCAGAGGAGTGAACCTCGC	780
DB	721	AGGAAGAGTGTGACGGGCGCGGAGAGAGAGAGCGCAGAGGAGTGAACCTCGC	780
QY	781	ACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCAATGAGCGCAACT	840
DB	781	ACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCAATGAGCGCAACT	840
QY	841	CCAAAGTGGCAGAGAAATGAGAGTGGTGGAGGCTCAAGAGCTGATTGAGAGT	900
DB	841	CCAAAGTGGCAGAGAAATGAGAGTGGTGGAGGCTCAAGAGCTGATTGAGAGT	900
QY	901	ATGAGCTGGCGAGGAGCATATGCAAAAGTCTTCAAACCAAGAGGCTTACACAGCAGC	960
DB	901	ATGAGCTGGCGAGGAGCATATGCAAAAGTCTTCAAACCAAGAGGCTTACACAGCAGC	960
QY	961	TGTTGGATGCAAGCTCCAGAGGCCAGAGATGCTTAAAGGAGGAGAGAGCGGACCC	1020
DB	961	TGTTGGATGCAAGCTCCAGAGGCCAGAGATGCTTAAAGGAGGAGAGAGCGGACCC	1020

DB	961	TGTTGGATGCAAGCTCCAGAGGCCAGAGATGCTTAAAGGAGGAGAGAGCGGACCC	1020
QY	1021	AGCGGAGAGAGGATTTCTCTGAAAGAGGAGTGAAGTCCAGAGGATGTGTGAGCTGA	1080
DB	1021	AGCGGAGAGAGGATTTCTCTGAAAGAGGAGTGAAGTCCAGAGGATGTGTGAGCTGA	1080
QY	1081	TGAAGCAGCAGAGACCCACCTGGAAGCAACAGCTTGCCTATACACAGAGTTTGAGG	1140
DB	1081	TGAAGCAGCAGAGACCCACCTGGAAGCAACAGCTTGCCTATACACAGAGTTTGAGG	1140
QY	1141	AGTTCCAGAACACATTTTCCAAAAGCAGAGGTATTCACACATTCAAGCAGAGATGG	1200
DB	1141	AGTTCCAGAACACATTTTCCAAAAGCAGAGGTATTCACACATTCAAGCAGAGATGG	1200
QY	1201	AAAAGATGATTAAGAGATCAAGAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGT	1260
DB	1201	AAAAGATGATTAAGAGATCAAGAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGT	1260
QY	1261	GGGAGAGCAGCAACCAAGGCCCTGTGTGATGGCTGAGGAGAAACAGTCCGGATTAAG	1320
DB	1261	GGGAGAGCAGCAACCAAGGCCCTGTGTGATGGCTGAGGAGAAACAGTCCGGATTAAG	1320
QY	1321	AACTGAGGGCTGTGAGGTAAATCCAAACGGCTGGAGAACTGTCCGGGATTAAG	1380
DB	1321	AACTGAGGGCTGTGAGGTAAATCCAAACGGCTGGAGAACTGTCCGGGATTAAG	1380
QY	1381	CAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTGTGGTCCAGGGCTCCC	1440
DB	1381	CAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTGTGGTCCAGGGCTCCC	1440
QY	1441	TCACGTGACAGTGGCCCTTGAGAGGAGCCAGAGGGGCTTGGGGCTCAAGCAGCCAGCTCC	1500
DB	1441	TCACGTGACAGTGGCCCTTGAGAGGAGCCAGAGGGGCTTGGGGCTCAAGCAGCCAGCTCC	1500
QY	1501	CCAGGTCACAGAGCGCTTGTCTACCCAGAGCACCGAGCACAGAGCATCAGGCCAGA	1560
DB	1501	CCAGGTCACAGAGCGCTTGTCTACCCAGAGCACCGAGCACAGAGCATCAGGCCAGA	1560
QY	1561	CTGGGCTCAAGAGCCACCTCCGGCAGGGCTTAGAGAGCTGTGTGGGTCTATGTGG	1620
DB	1561	CTGGGCTCAAGAGCCACCTCCGGCAGGGCTTAGAGAGCTGTGTGGGTCTATGTGG	1620
QY	1621	GAAGGAGCGGAGCCAGCGCTGGCCCATAAAGGCTCCCATCTGAGAGAGCCCA	1680
DB	1621	GAAGGAGCGGAGCCAGCGCTGGCCCATAAAGGCTCCCATCTGAGAGAGCCCA	1680
QY	1681	TTGCTGAAGCCAGAGTGTCTTGACCTGGCTGGCATCTGGCACTTGCATTTTGGATTTT	1740
DB	1681	TTGCTGAAGCCAGAGTGTCTTGACCTGGCTGGCATCTGGCACTTGCATTTTGGATTTT	1740
QY	1741	GTGGTCAAGTTTACGTACATAGGGCAATTTTGAAGGCTTGCATTTTGAATTTTGA	1800
DB	1741	GTGGTCAAGTTTACGTACATAGGGCAATTTTGAAGGCTTGCATTTTGAATTTTGA	1800
QY	1801	TAAGTGTACAGTGGCTTGCATTTGGGATGGGGTGTACAGATGAAGTCAGTGGCTTG	1860
DB	1801	TAAGTGTACAGTGGCTTGCATTTGGGATGGGGTGTACAGATGAAGTCAGTGGCTTG	1860
QY	1861	TCGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCTATCTGTAGCTGCCATCAGATGAGTTG	1920
DB	1861	TCGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCTATCTGTAGCTGCCATCAGATGAGTTG	1920
QY	1921	GCAGAGTGTACCTGAGCATTTTCTGTCTGATTTGAGGCTCAGAGCCCTCCCTGCCCTTT	1980
DB	1921	GCAGAGTGTACCTGAGCATTTTCTGTCTGATTTGAGGCTCAGAGCCCTCCCTGCCCTTT	1980
QY	1981	CAGAGCTCAAAACCAAGTAATACACCAAGTCTTGACCTGCAATTTCTTGTGAGCAGGCT	2040
DB	1981	CAGAGCTCAAAACCAAGTAATACACCAAGTCTTGACCTGCAATTTCTTGTGAGCAGGCT	2040
QY	2041	TGCTTGGTCAAGTCCAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCT	2100
DB	2041	TGCTTGGTCAAGTCCAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCT	2100


```
Db 2989 CTTCTCTCTGTGGGACCCATAGCTGTCTCTAAGGCTAGAAATGAATGAACCTA 3048
QY 2877 TTGACTGGGCCCCAGAAATCATAAATGGCTGAGACAGTGTCTGTGCTCTGTCT 2936
Db 3049 TTGACTGGGCCCCAGAAATCATAAATGGCTGAGACAGTGTCTGTGCTCTGTCT 3108
QY 2937 ACCCCCACTCCAGTACATACTACTATGTACTGTGTAGAGCCATCTATATGCTGAATGT 2996
Db 3109 ACCCCCACTCCAGTACATACTACTATGTACTGTGTAGAGCCATCTATATGCTGAATGT 3168
QY 2997 TCTGCTGTGCAAACTTGCAGGGTATAGCCAGTGTGTGCAAGCAGTTTTCGGGGA 3056
Db 3169 TCTGCTGTGCAAACTTGCAGGGTATAGCCAGTGTGTGCAAGCAGTTTTCGGGGA 3228
QY 3057 CAACAGATGACTCAGACCAAGATGAGATGATGTTAGGGCTTGTCTTCTGCTGTTTT 3116
Db 3229 CAACAGATGACTCAGACCAAGATGAGATGATGTTAGGGCTTGTCTTCTGCTGTTTT 3288
QY 3117 TCTTTGAA-CTAG-TCATTTGCTGCGAGTCCCTTCACTTCCATACCTAGCCCACTTT 3174
Db 3289 TCTTTGAGCTAGTTCATTTGCTGCGAGTCCCTTCACTTCCATACCTAGCCCACTTT 3348
QY 3175 TTAGCCCTTACCTTAATCTCTCAGATAAGTGTGTTTCCACAAAGATGTTAAGTACTGAAT 3234
Db 3349 TTAGCCCTTACCTTAATCTCTCAGATAAGTGTGTTTCCACAAAGATGTTAAGTACTGAAT 3408
QY 3235 CATGTGTGACTGAGACAGAGATGGCAATGAATGGCACACCAATTTCTCTCTGCTGCC 3294
Db 3409 CATGTGTGACTGAGACAGAGATGGCAATGAATGGCACACCAATTTCTCTCTGCTGCC 3468
QY 3295 CCAGGGCAGGTACCACTGATCTGATCAGAGTGGCTGCTATCTCTGCTGATCTCTCA 3354
Db 3469 CCAGGGCAGGTACCACTGATCTGATCAGAGTGGCTGCTATCTCTGCTGATCTCTCA 3528
QY 3355 CATCTAGTGTCCCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGA 3414
Db 3529 CATCTAGTGTCCCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGA 3588
QY 3415 TACTGTCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTT 3461
Db 3589 TACTGTCTCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTT 3635

RESULT 7
HSM803646 4851 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp451K215 (from clone DKFZp451K215);
DEFINITION complete cds.
ACCESSION AL832338
VERSION AL832338.1 GI:21732900
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4851)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
Direct Submission
TITLE Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BWPZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp451K215) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source
Location/Qualifiers
1..4851
```

```
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="RZPD:DKFZp451K215"
/db_xref="taxon:9606"
/map="lp34.2-36.11"
/clone="DKFZp451K215"
/tissue_type="human skeletal muscle"
/clone_lib="451 (synonym: hicc1). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4851
/gene="DKFZp451K215"
103..1743
/gene="DKFZp451K215"
/codon_start=1
/protein="hypothetical protein"
/protein_id="CAD91138.1"
/db_xref="GI:30722346"
/translation="MKNQDKNGAAKOSNPKSPQPEAGPEGAQERPSQAAPAVEAE
GFGSQAPRKPEGAQARTAQALRDVSEELSRQLEDILSTYCVNNQGGPGEAGQ
EPAPEDAEKSRVVARNGEPEPTPVVNGEPEKSGDPNTEIRQSDEVDGRHRRPQ
EKKAAGIGKBITILLMOTLNTLSTPEEKLAALCKYAELEHRSNOKMKLLKOKOS
OLVQEKHLRGEHSKAVLRSKLSLELQELRHNSLKEGVQARSEERKKEVTS
FQVTLNDLQMEQHNSRNSKLRNVELAERLKLIEQELREHIDKVFKKDLQO
QVLDAXUQAQAEMLKAEERHREKDFLLKEAVESQRMCELMKQOETHLKKQALYTE
KFEFQNTLSKSEVFTTFKQEMKMTTKIKLEKETTMYRSRWESNKALLEAEK
TVRKELRGLQVKIQRLEKLCALQTERNDLNKRVQDLISAGGQSLTDSGPERPEP
GAQAPSSPRVTEAPCYGAPSAEASGQTGPQPTSARA"
4814..4819
/polya_signal
/gene="DKFZp451K215"
4834
/polya_site
/gene="DKFZp451K215"
4834
ORIGIN
Query Match 39.1%; Score 1837; DB 9; Length 4851;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3457; Conservative 0; Mismatches 12; Indels 9; Gaps 8;
QY 1 CAAAGAAGCAGCCAGGACAAACCGAAGCAGAGCCGAGGGAGCCAGAGCGGCCAGCC 60
Db 149 CAAAGAAGCAGCCAGGACAAACCGAAGCAGAGCCGAGGGAGCCAGAGCGGCCAGCC 208
QY 61 AGGCGGCTCTGCGAGTAGAAGCAGAGGTCCTCCGCGAGCAGCAGGCTCTCGGAAGCGG 120
Db 209 AGGCGGCTCTGCGAGTAGAAGCAGAGGTCCTCCGCGAGCAGCAGGCTCTCGGAAGCGG 268
QY 121 AGGCGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 180
Db 269 AGGCGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 328
QY 181 GCGGCGAATCGAAGACATCTGAGCAGCATACTGTGGCAATAAACACAGGGGGGCCCCG 240
Db 329 GCGGCGAATCGAAGACATCTGAGCAGCATACTGTGGCAATAAACACAGGGGGGCCCCG 388
QY 241 GCGAGATGGGGGCACAGGGTGAAGCGGCTGAACCGGCTGAACCGGAGATGCGAGAAGTCCCGGACCT 300
Db 389 GCGAGATGGGGGCACAGGGTGAAGCGGCTGAACCGGCTGAACCGGAGATGCGAGAAGTCCCGGACCT 448
QY 301 ATGTGGCAAGAAATGGGGAGCCTGAACCACTCAGTAGTCTATGGAGAGAGAGAACCT 360
Db 449 ATGTGGCAAGAAATGGGGAGCCTGAACCACTCAGTAGTCTATGGAGAGAGAGAACCT 508
QY 361 CCAAGGGGATTCACAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGAGCCGAGACC 420
Db 509 CCAAGGGGATTCACAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGAGCCGAGACC 568
QY 421 ATCGAAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAGTTGCTGA 480
Db 569 ATCGAAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAGTTGCTGA 628
QY 481 TGCAGACATTGAATCTCTGTAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 540
Db 629 TGCAGACATTGAATCTCTGTAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGT 688
```


Db	2847	CAGTGGCTCTTGGCTTCATGGGTTTTCTGCTCTCTGCCCCCACCACACATGTCAC	2906
Qy	2757	TCCTCAAGATTGTCCTGATTTATTTCTCGGCACCTCCCTGCTGCTCTGCGGATTC	2816
Db	2907	TCCTCAAGATTGTCCTGATTTATTTCTCGGCACCTCCCTGCTGCTCTGCGGATTC	2966
Qy	2817	ACTTCTTCTGCTGGG-GCCATAGCTGTGTCTACAGTAAGAATGAATGAACT	2875
Db	2967	ACTTCTTCTGCTGGGAGCCATAGCTGTGTCTACAGTAAGAATGAATGAACT	3026
Qy	2876	ATTGACTGGGCCCCAGAAATCCATAAATGGCTCGACAGAGTTGTTCTGTCTCTGTT	2935
Db	3027	ATTGACTGGGCCCCAGAAATCCATAAATGGCTCGACAGAGTTGTTCTGTCTCTGTT	3086
Qy	2936	TACCCCACTCCAGTACATACTACTATGTCTGTGTAGAGCCATTTCTATGCTGAATG	2995
Db	3087	TACCCCACTCCAGTACATACTACTATGTCTGTGTAGAGCCATTTCTATGCTGAATG	3146
Qy	2996	TTCTGCTGTGCAAACTTGGCAGGTTATAGCCAGTGTGTTGTCACAGCAGTTTTCGGG	3055
Db	3147	TTCTGCTGTGCAAACTTGGCAGGTTATAGCCAGTGTGTTGTCACAGCAGTTTTCGGG	3206
Qy	3056	ACAACAGAACTGACACAGCAAGATGATAGGATGGTTAGGGCTTTCCTTCTGCTGTT	3115
Db	3207	ACAACAGAACTGACACAGCAAGATGATAGGATGGTTAGGGCTTTCCTTCTGCTGTT	3266
Qy	3116	TTCTTTGAA-CTAG-TCATTGCTCTGAGGCTCCCTTCATCTTCATACCTAGCCACTCT	3173
Db	3267	TTCTTTGAACTGATTCATTGCTCTGAGGCTCCCTTCATCTTCATACCTAGCCACTCT	3326
Qy	3174	TTTAGCCCTTACCTTAATCTCTCAGATAAGTTGGTTTCAAAAAGATGTTAAGTACTGAA	3233
Db	3327	TTTAGCCCTTACCTTAATCTCTCAGATAAGTTGGTTTCAAAAAGATGTTAAGTACTGAA	3386
Qy	3234	TCATGTGTGACTGACAGAGATGGCAATGAATGACACACATTTCTCTCTCTGTC	3293
Db	3387	TCATGTGTGACTGACAGAGATGGCAATGAATGACACACATTTCTCTCTCTGTC	3446
Qy	3294	CCAGGGCAGGTACCATGATCTGCATCAGATGGCTGCTGATTTCTCTGCTGATCTTC	3353
Db	3447	CCAGGGCAGGTACCATGATCTGCATCAGATGGCTGCTGATTTCTCTGCTGATCTTC	3506
Qy	3354	ACATCTAGGTGCTCTAAGCAGCTGTGTGAGTGTGAGATCTCTGCAATCTCTGCTGAG	3413
Db	3507	ACATCTAGGTGCTCTAAGCAGCTGTGTGAGTGTGAGATCTCTGCAATCTCTGCTGAG	3566
Qy	3414	ATACTGCTGCTGCTGAGTGTGCTCCATGACCTTTTCTTCCCTTTGAACTCCCTCT	3471
Db	3567	ATACTGCTGCTGCTGAGTGTGCTCCATGACCTTTTCTTCCCTTTGAACTCCCTCT	3624
RESULT 8			
AF516206		1993 bp mRNA linear	PRI 15-FEB-2003
LOCUS			
DEFINITION		Homo sapiens taxilin mRNA, complete cds.	
ACCESSION		AF516206	
VERSION		AF516206.1	GI:28394223
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE			
AUTHORS		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 1993)	
JOURNAL		Nogami, S., Satoh, S., Nakano, M., Shimizu, H., Fukushima, H., Maruyama, A., Terano, A. and Shirataki, H.	
MEDLINE		Taxilin; a novel syntaxin-binding protein that is involved in Ca2+-dependent exocytosis in neuroendocrine cells	
PubMed		Genes Cells 8 (1), 17-28 (2003)	
REFERENCE		2 (bases 1 to 1993)	
AUTHORS		Shirataki, H., Nogami, S. and Satoh, S.	

TITLE	Direct Submission
JOURNAL	Submitted (29-MAY-2002) Division of Molecular and Cell Biology, Institute for Medical Sciences, Dokkyo University School of Medicine, 880 Kitakobayashi, Mibu-machi, Tochigi 321-0293, Japan
FEATURES	Location/Qualifiers
source	1..1993 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="1" /map="p36.11-p34.2" /tissue_type="brain" 79..1719 /codon_start=1 /product="taxilin" /protein_id="AA042465.1" /db_xref="GI:28394224" /translation="MKNDKXGAAKQSNPKSPGQPEAPGEPGAQERPSQAAPAVEAE GPSSQAPKPEGAQARTAQSGALRDVSELSROLEDILSTYCVNNQGGPDGDAQ EPABPEDAKSRITVVRNGEPEPTVNGEKEPSKDPNTEIIRSDVGGDRHRPQ EKKAKGLGKSTITLLMOTLNTLSPEKIALCKYAELEEHNSOKMKLKKOSQ QIVQKDLRHSKAVLARSLSLRELORHNSLKEGVQARSEEEKKRVTSK FQVTLNDLOQMEQHNRNSKLQENMELARLKLQYELREHIDKVPKHLQJQ QLVDAKQQAQEMLEAEERHOREKDFLLKEAVESQRMELMQOETHLKKQALYTE KFEFQNTLSKSEVFTFKQEMKMTKKIKLEKETTMYRSWESNKALLEAEK TVRKELEGLOVKIQRLEKLCRALQTERNDLNRVQDLSAGGQGLSDSGPERRPEG GAQAPSSPRVTEAPCYPGAPSTASGQTGPQPTSARA"
CDS	
ORIGIN	
Query Match	36.2%; Score 1698; DB 9; Length 1993;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1868; Conservative	0; Mismatches 1; Indels 1; Gaps 1;
Qy	1 CAAAGCAGCCAGGACCAACCGGAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCC 60
Db	125 CAAAAGCAGCCAGGACCAACCGGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCC 184
Qy	61 AGCGGCTCTCTGAGTAGAACGAGAAGGTCCCGGACGAGCCAGGCTCTCTCGAAGCGCG 120
Db	185 AGCGGCTCTCTGAGTAGAACGAGAAGGTCCCGGACGAGCCAGGCTCTCTCGAAGCGCG 244
Qy	121 AGGGGCTCAAGCGAAGCGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGTGAGGAGCTGA 180
Db	245 AGGGGCTCAAGCGAAGCGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGTGAGGAGCTGA 304
Qy	181 GCGCCAACTGGAAGACATCTGAGCACATCTGTTGACAAATAACACGGGGGGGCCCG 240
Db	305 GCGCCAACTGGAAGACATCTGAGCACATCTGTTGACAAATAACACGGGGGGGCCCG 364
Qy	241 GCGAGATGGGCGACAGGTCGAGCGGCTGAACCCGAAAGATGCGAGAGATGTCGAGACCT 300
Db	365 GCGAGATGGGCGACAGGTCGAGCGGCTGAACCCGAAAGATGCGAGAGATGTCGAGACCT 424
Qy	301 ATGTGCAAGGAATGGGAGCCTGAACCAACTCCAGTACTGATGAGAGAGAGAACCT 360
Db	425 ATGTGCAAGGAATGGGAGCCTGAAACCACTCCAGTACTCAATGGAGAGAGAACCT 484
Qy	361 CCAAGGGGATCCAAACACAGAGAGATCCCGGAGAGTGCAGGTCGAGACCGAGACC 420
Db	485 CCAAGGGGATCCAAACACAGAGAGATCCCGGAGAGTGCAGGTCGAGACCGAGACC 544
Qy	421 ATCGAGGCCACAGGAGAGAAAAGCAAGGTTTGGGAGGAGATCAGTTGCTGA 480
Db	545 ATCGAGGCCACAGGAGAGAAAAGCAAGGTTTGGGAGGAGATCAGTTGCTGA 604
Qy	481 TGCAGACATTGAATCTCTGAGTACCCCGAGGAGAGCTGGCTGCTCTGTGCAAGAAGT 540
Db	605 TGCAGACATTGAATCTCTGAGTACCCCGAGGAGAGCTGGCTGCTCTGTGCAAGAAGT 664
Qy	541 ATGCTGAATCTGTGAGGAGCAGCGGATTCACAGAGCAGATGAAGCTCTCTACAGAAA 600
Db	665 ATGCTGAATCTGTGAGGAGCAGCGGATTCACAGAGCAGATGAAGCTCTCTACAGAAA 724

1981	Qy	1981	CAGAGCTCAAAACAAAGTAATACACCAAGGCTCTGACATGCAATTTGTCTTTGTGACAGAGGCT	2040
2306	Db	2306	CAGAGCTCAAGACAAGTAATACACCAAGGCTCTGACATGCAATTTGTCTTTGTGACAGAGGCT	2365
2041	Qy	2041	TGCTTTGGTCAAGCTCAGGCCCTCCTAGCTGCT-TGGAGGCTCCCTTTGATTTCTCTAGACCTG	2099
2366	Db	2366	TGCTTTGGTCAAGCTCAGGCCCTCCTAGCTGCTCTGGAGGCTCCCTTTGATTTCTCTAGACCTG	2425
2100	Qy	2100	GAAGAAGTGTCCTTAGCAGAGGCCCTGGCAGGGCGCTCAGAGCT-GGGATTTCTCTGCCTG	2158
2426	Db	2426	GAAGAAGTGTCCTTAGCAGAGGCCCTGGCAGGGCGCTCAGAGCTGGGGATTTCTCTGCCTG	2485
2159	Qy	2159	GAACAAGGACCTGGGAATGTTTTTCGGTGGGATGATGTGCTGTCAGAGAGCCCTTGG	2218
2486	Db	2486	GAACAAGGACCTGGGAATGTTTTTCGGTGGGATGATGTGCTGTCAGAGAGCCCTTGG	2545
2219	Qy	2219	GCATCGCTTCCCTCCCTTTGGTAGTGCCAGGACCAAGGCCAATGATGCTTCTCAGTAGC	2278
2546	Db	2546	GCATCGCTTCCCTCCCTTTGGTAGTGCCAGGACCAAGGCCAATGATGCTTCTCAGTAGC	2605
2279	Qy	2279	CTTATCATTTACAGGTGCTCTTAGCCTGCAAAATGATGTACAAGAGATCACCNAAG	2338
2606	Db	2606	CTTATCATTTACAGGTGCTCTTAGCCTGCAAAATGATGTACAAGAGATCACCNAAG	2665
2339	Qy	2339	GATTAATTTCTGAAGTG-TTTTTTCTTTAATTTCTTTTTC--TTTTTTTTTTTTTTCTTTT	2395
2666	Db	2666	GATTAATTTCTGAAGTGTTTTTTCTTTAATTTCTTTTCTTTTCTTTTTTTTTTTCTTTT	2725
2396	Qy	2396	TCTTTTTTTTTGCATATGACAGTGTTTGATTTAGGACCTTCCAGGAAAAAGGATGCT	2455
2726	Db	2726	TCTTTTTTTTTGCATATGACAGTGTTTGATTTAGGACCTTCCAGGAAAAAGGATGCT	2785
2456	Qy	2456	GTACCACTGGTGCTCGGTGCTGGCTCCAGTGTGCCACTCCTCTCACACACCCACTTG	2515
2786	Db	2786	GTACCACTGGTGCTCGGTGCTGGCTCCAGTGTGCCACTCCTCTCACACACCCACTTG	2845
2516	Qy	2516	GCTCCTTTGCCAATCTGATGCTGAGTTTCTGTCTTGGTGAGATCAGTTGTTTGTGGTA	2575
2846	Db	2846	GCTCCTTTGCCAATCTGATGCTGAGTTTCTGTCTTGGTGAGATCAGTTGTTTGTGGTA	2905
2576	Qy	2576	AAAGAAAGGAAGGCTCTTGATGGCTTGGCCACAAGCTTACCTGTGGTTCAGTCTCTG	2635
2906	Db	2906	AAAGAAAGGAAGGCTCTTGATGGCTTGGCCACAAGCTTACCTGTGGTTCAGTCTCTG	2965
2636	Qy	2636	AGAGGCCACCAACAGTTCCCATCAGACTGTCTCCATCGACAGTGTCTCGGTCCCATGT	2695
2966	Db	2966	AGAGGCCACCAACAGTTCCCATCAGACTGTCTCCATCGACAGTGTCTCGGTCCCATGT	3025
2696	Qy	2696	CCAGCTGCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCCCCCACCCCAATGTGCA	2755
3026	Db	3026	CCAGCTGCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCCCCCACCCCAATGTGCA	3085
2756	Qy	2756	ATCCTCAAGATTTCTCTGATTTCTAATTCCTGGCACCTCCTGCTGCTTGGGGATTC	2815
3086	Db	3086	ATCCTCAAGATTTCTCTGATTTCTAATTCCTGGCACCTCCTGCTGCTTGGGGATTC	3145
2816	Qy	2816	TACTTCTTCTGCTGGG- GCCCATAGCTGTGTCTAACAGGTGAAGAAATGAATGAAC	2874
3146	Db	3146	TACTTCTTCTGCTGGGAGCCATAGCTGTGTCTAACAGGTGAAGAAATGAATGAAC	3205
2875	Qy	2875	TATTGACTGGGCCCCAGAAATCCATAAAATGGCTGCAGAGTTGTTTTCTGTGTCCTGTT	2934
3206	Db	3206	TATTGACTGGGCCCCAGAAATCCATAAAATGGCTGCAGAGTTGTTTTCTGTGTCCTGTT	3265
2935	Qy	2935	CTACCCCACTCCAGTACATAACTACTATGTACTGTGTAGAGCCATTCTATGTCTGAAT	2994
3266	Db	3266	CTACCCCACTCCAGTACATAACTACTATGTACTGTGTAGAGCCATTCTATGTCTGAAT	3325
2995	Qy	2995	GTTCTGCTGTTGAAAATTGCCAGGTTATAGCCAGTGTTTGTGCCAAGCAGTTTCCGG	3054
3326	Db	3326	GTTCTGCTGTTGAAAATTGCCAGGTTATAGCCAGTGTTTGTGCCAAGCAGTTTCTCG	3385
3055	Qy	3055	GACAACAGAAATGACTCAGACAAGATGGATAGGATGGTTAGGGCTTTGCTTCTGCTGTT	3114

Db	3386	GACAAACAGATGACTCAGACCAAGATGGATAGGATGGTTAGGGCTTTGCTTCTTGCTGTT	3445
Qy	3115	TTTCTTTGAA	3124
Db	3446	TTTCTTTGAA	3455
RESULT 10			
AR409338			
LOCUS	AR409338	1638 bp	DNA linear PAT 18-DEC-2003
DEFINITION	Sequence 46 from patent US 6632923.		
ACCESSION	AR409338		
VERSION	AR409338.1	GI:40160126	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1638)		
AUTHORS	Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.		
TITLE	Low density lipoprotein binding proteins and their use in		
JOURNAL	diagnosing and treating atherosclerosis		
FEATURES	Patent: US 6632923-A 46 14-OCT-2003;		
source	Location/Qualifiers		
	1..1638		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	30.6%;	Score 1439;	DB 6; Length 1638;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1589;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;
Qy	1	CAAAAACAGCCCGCAGGACAACCGGAAGCAGGACCCGAGGGAGCCCGCAGGAGCGGCCCGCAGCC	60
Db	47	CAAAAACAGCCCGCAGGACAACCGGAAGCAGGACCCGAGGGAGCCCGCAGGAGCGGCCCGCAGCC	106
Qy	61	AGCGCGCTCTGAGTAGAGACGAGAGGTCCCGCAGCAGCCAGGCTCTCTCGGAAGCGG	120
Db	107	AGCGCGCTCTGAGTAGAGACGAGAGGTCCCGCAGCAGCCAGGCTCTCTCGGAAGCGG	166
Qy	121	AGGGGGCTCAAGCCAGAACCGCTCAGTCTCGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA	180
Db	167	AGGTGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGA	226
Qy	181	GCGCCAACTGGAGACATATGAGACATATCTGTGTGGACATATACAGGGGGGCCCG	240
Db	227	GCGCCAACTGGAGACATATGAGACATATCTGTGTGGACATATACAGGGGGGCCCG	286
Qy	241	GCAGGAGTGGGCACAGGGTGAGCGGCTCAACCCGGAAGATGCAGAGAAAGTCCCGGACCT	300
Db	287	GCAGGAGTGGGCACAGGGTGAGCGGCTCAACCCGGAAGATGCAGAGAAAGTCCCGGACCT	346
Qy	301	ATGTGGCAAGGANTGGGAGCGCTGAAACCACTCCAGTAGTCTATGGAGAGAGAAACCT	360
Db	347	ATGTGGCAAGGANTGGGAGCGCTGAAACCACTCCAGTAGTCAATGGAGAGAGAAACCT	406
Qy	361	CCAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCCGAGACC	420
Db	407	CCAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCCGAGACC	466
Qy	421	ATCGAAGGCCACAGGAGAAGAAAAAGCCAAAGGGTTTGGGGAAGGAGATCACTGTGCTGA	480
Db	467	ATCGAAGGCCACAGGAGAAGAAAAAGCCAAAGGGTTTGGGTAAGGAGATCACGTTGCTGA	526
Qy	481	TGCAGACATTGATCTCTGAGTACCCACAGAGGAGAGCTGGCTGCTCTGTGCAAGAGT	540
Db	527	TGCAGACATTGATCTCTGAGTACCCACAGAGGAGAGCTGGCTGCTCTGTGCAAGAGT	586
Qy	541	ATGCTGAACCTGCTGGAGAGCACCGGAATTCACAGAGACGATGAAGCTCTTACAGAAA	600
Db	587	ATGCTGAACCTGCTGGAGAGCACCGGAATTCACAGAGACGATGAAGCTCTTACAGAAA	646

601 AGCAGAGCAGCTGGTCAAGAGAGAGACACCTGCGCGGTGAGCAGCAGCAAGCGCGTCC 660
Db AGCAGAGCAGCTGGTCAAGAGAGAGACACCTGCGCGGTGAGCAGCAGCAAGCGCGTCC 706
661 TGGCCCGCAGCAAGCTTGGAGAGCCTATGCGGTGAGCTGCGAGCGCAACAACGCTTCCCTCA 720
Db TGGCCCGCAGCAAGCTTGGAGAGCCTATGCGGTGAGCTGCGAGCGCAACAACGCTTCCCTCA 766
721 AGGAGAGAGTGTCCAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGAGTGCACCTCCG 780
Db AGGAGAGAGTGTCCAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGAGTGCACCTCCG 826
767 AGGAGAGAGTGTCCAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGAGTGCACCTCCG 840
Db AGGAGAGAGTGTCCAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGAGTGCACCTCCG 886
781 ACTTCCAGGTGACACTGAATGACATTCAGCTGCGAGTGGAAACGCAAAATGAGCGCAACT 900
Db ACTTCCAGGTGACACTGAATGACATTCAGCTGCGAGTGGAAACGCAAAATGAGCGCAACT 946
841 CCAAGCTGCGCAAGAGACATGAGAGCTGCTGAGAGCTCAAGAGCTGATTTGAGCAGT 960
Db CCAAGCTGCGCAAGAGACATGAGAGCTGCTGAGAGCTCAAGAGCTGATTTGAGCAGT 1006
901 ATGAGCTGCGGAGGAGCATATGACAAAGTCTTCAACACAGAGCTTCAACAGCAGCAGC 1020
Db ATGAGCTGCGGAGGAGCATATGACAAAGTCTTCAACACAGAGCTTCAACAGCAGCAGC 1066
961 TGGTGGATGCCAAGCTCCAGAGCCCGCAGAGAGTCTTAAAGAGGAGCAGAGAGCGGCACC 1080
Db TGGTGGATGCCAAGCTCCAGAGCCCGCAGAGAGTCTTAAAGAGGAGCAGAGAGCGGCACC 1126
1021 AGCGGAGAGAGGATTTCTCTGAGAGAGGAGGAGGAGTCCAGAGAGTGTGAGCTGA 1140
Db AGCGGAGAGAGGATTTCTCTGAGAGAGGAGGAGGAGTCCAGAGAGTGTGAGCTGA 1186
1081 TGAAGCAGCAGAGAGACCCACCTGAGCAACAGCTTGGCTTATACAGAGAGATTTGAGG 1200
Db TGAAGCAGCAGAGAGACCCACCTGAGCAACAGCTTGGCTTATACAGAGAGATTTGAGG 1246
1201 AAAAGATGACTAAGAGATCAAGAGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db AAAAGATGACTAAGAGATCAAGAGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1306
1261 GGGAGAGCAGCAACAGCGCTGAGAGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db GGGAGAGCAGCAACAGCGCTGAGAGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1366
1321 AACTGGAGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Db AACTGGAGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1426
1381 CAGAGCGCAATGACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db CAGAGCGCAATGACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1486
1441 TCAGTGCAGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
Db TCAGTGCAGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1546
1501 CCAGGCTCAGAGAGCGCTTGTACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db CCAGGCTCAGAGAGCGCTTGTACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1606
1561 CTGGGCTCAGAGAGCGCTTGTACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592
Db CTGGGCTCAGAGAGCGCTTGTACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1638

RESULT 11
AX239603 1638 bp DNA linear PAT 26-SEP-2001
LOCUS
DEFINITION Sequence 46 from Patent WO0164874.

ACCESSION AX239603
VERSION AX239603.1 GI:15797277
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE Low density lipoprotein binding proteins and their use in diagnosing
JOURNAL and treating atherosclerosis
PATENT WO 0164874-A 46 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
source
CDS
ORIGIN

Query Match 30.6%; Score 1439; DB 6; Length 1638;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 60
Db 47 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 106
QY 61 AGCGGCTCTTCAGTACAGCAGAGGAGTCCCGGAGCAGCAGGAGTCTCTCGGAGCCGG 120
Db 107 AGCGGCTCTTCAGTACAGCAGAGGAGTCCCGGAGCAGCAGGAGTCTCTCGGAGCCGG 166
QY 121 AGCGGCTCTTCAGTACAGCAGAGGAGTCTCTCGGAGCCCTTCGTGATGTCTCTGAGGAGTGA 180
Db 167 AGCGGCTCTTCAGTACAGCAGAGGAGTCTCTCGGAGCCCTTCGTGATGTCTCTGAGGAGTGA 226
QY 181 GCGGCAACTGGAAGACATCTGAGCAGCAGTCTCTCGGAGCAGCAGGAGTCTCTCGGAGCCGG 240
Db 227 GCGGCAACTGGAAGACATCTGAGCAGCAGTCTCTCGGAGCAGCAGGAGTCTCTCGGAGCCGG 286
QY 241 GCGGAGTGGGACACAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 300
Db 287 GCGGAGTGGGACACAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 346
QY 301 ATGTGGCAAGGAGTGGGAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 360
Db 347 ATGTGGCAAGGAGTGGGAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 406
QY 361 CCAAGGAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 420
Db 407 CCAAGGAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 466
QY 421 ATCGAAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 480
Db 467 ATCGAAGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCC 526
QY 481 TCGCAGCATTTGAGTCTCTGAGTACCCAGGAGGAGGAGTGGTCTCTCTGTCAGGAGT 540

Db 527 TGCAGACATTGAATCTCTCAGTACCCACAGAGAGAGCTGGCTGCTGTGCAAGAGT 586
QY 541 ATGCTGAACCTGCTGGAGGACACCGAATTCACAGAGCAGATGAAGCTCTCTACAGAAA 600
Db 587 ATGCTGAACCTGCTGGAGGACACCGAATTCACAGAGCAGATGAAGCTCTCTACAGAAA 646
QY 601 AGCAGAGCCAGCTGGTGCAGAGAGAGACACCTGCGCGGTGAGCAGACAGAGCCGCTCC 660
Db 647 AGCAGAGCCAGCTGGTGCAGAGAGAGACACACCTGCGCGGTGAGCAGACAGAGCCGCTCC 706
QY 661 TGGCCCGCAGACAGCTTGAGAGCCTATCCGCTGAGCTGAGCGGCACACCCGCTCCCTCA 720
Db 707 TGGCCCGCAGACAGCTTGAGAGCCTATCCGCTGAGCTGAGCGGCACACCCGCTCCCTCA 766
QY 721 AGGAAGAAGCTGTGCAGCGCGCCGCGAGGAGGAGAGAGCCAGAGAGGTGACCTCGC 780
Db 767 AGGAAGAAGCTGTGCAGCGCGCCGCGAGGAGGAGAGAGCCAGAGAGGTGACCTCGC 826
QY 781 ACTTCAGAGTGAACCTGAATTCAGCTGAGATGGACAGACATGAGCGCACT 840
Db 827 ACTTCAGAGTGAACCTGAATTCAGCTGAGATGGACAGACATGAGCGCACT 886
QY 841 CCAAGCTGCGCCCAAGAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 900
Db 887 CCAAGCTGCGCCCAAGAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 946
QY 901 ATGAGTGGCGAGGAGCATATCGACAAAGCTTTCAAAACACAGAGACCTACACAGCAGC 960
Db 947 ATGAGTGGCGAGGAGCATATCGACAAAGCTTTCAAAACACAGAGACCTACACAGCAGC 1006
QY 961 TGGTGATGCGAAGCTCCAGCGGCCCGAGGAGATGCTTAAGGAGGAGAGAGCGGCACC 1020
Db 1007 TGGTGATGCGAAGCTCCAGCGGCCCGAGGAGATGCTTAAGGAGGAGAGAGCGGCACC 1066
QY 1021 AGCGGGAGAGAGATTTCTCTCTGAAGAGGAGCTAGAGTCCAGAGAGATGTGTGAGCTGA 1080
Db 1067 AGCGGGAGAGAGATTTCTCTCTGAAGAGGAGCTAGAGTCCAGAGAGATGTGTGAGCTGA 1126
QY 1081 TGAAGCAGAGAGAGACCCAGCTGAAGCAACAGCTTGCCTCTATACAGAGAGATTTGAGG 1140
Db 1127 TGAAGCAGAGAGAGACCCAGCTGAAGCAACAGCTTGCCTCTATACAGAGAGATTTGAGG 1186
QY 1141 AGTTCCAGAACACACTTTCAAAAGCAGCGAGGTATTCACCACTTCAAGCAGGAGATGG 1200
Db 1187 AGTTCCAGAACACACTTTCAAAAGCAGCGAGGTATTCACCACTTCAAGCAGGAGATGG 1246
QY 1201 AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAGAGAGAACCACTATGTACCGGTCCCGGT 1260
Db 1247 AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAGAGAGAACCACTATGTACCGGTCCCGGT 1306
QY 1261 GGGAGAGCAGCAACAGGCGCTGCTGAGATGCTGAGGAGAGAGAACAGTCCCGGATTAAG 1320
Db 1307 GGGAGAGCAGCAACAGGCGCTGCTGAGATGCTGAGGAGAGAGAACAGTCCCGGATTAAG 1366
QY 1321 AACTGGAGGCGCTGCGAGTAAAAATCCACGGCTGGAGAGCTGTGCGCGGCACTGACAG 1380
Db 1367 AACTGGAGGCGCTGCGAGTAAAAATCCACGGCTGGAGAGCTGTGCGCGGCACTGACAG 1426
QY 1381 CAGAGCGCAATGACTGAACAGAGGTACAGACCTGAGTGTGTGTGCGCAGGCGTCCC 1440
Db 1427 CAGAGCGCAATGACTGAACAGAGGTACAGACCTGAGTGTGTGTGCGCAGGCGTCCC 1486
QY 1441 TCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGCGCTGGGGCTCAAGCAACCCAGCTCCC 1500
Db 1487 TCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGCGCTGGGGCTCAAGCAACCCAGCTCCC 1546
QY 1501 CCAAGGTGACAGAGCGCTTGTGATCCAGAGAGCAGCGAGCAGAGAGATCAAGGCGCAG 1560
Db 1547 CCAAGGTGACAGAGCGCTTGTGATCCAGAGAGCAGCGAGCAGAGAGATCAAGGCGCAG 1606
QY 1561 CTGGCGCTCAAGAGCGCCACCTCCGCGAGGCGC 1592

Db 1607 CTGGCGCTCAAGAGCCACCTCCGCCAGGGCC 1638
RESULT 12
AX578043/c
LOCUS AX578043
DEFINITION Sequence 165 from Patent WO02081745.
ACCESSION AX578043
VERSION AX578043.1 GI:27647251
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J.,
Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.
TITLE Genes involved in osteogenesis, and methods of use
JOURNAL Patent: WO 02081745-A 165 17-OCT-2002;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..1793
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Homo sapiens interleukin 14 (IL14), mRNA"
ORIGIN
Query Match 27.5%; Score 1294; DB 6; Length 1793;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 585 GGAATTCACAGAGCAGATGAAGCTCTACAGAAAAGCAGACGACGCTGTGCAAGAGA 624
Db 1793 GGAATTCACAGAGCAGATGAAGCTCTACAGAAAAGCAGACGACGCTGTGCAAGAGA 1734
QY 625 AGGACCACTGCGCGTGCAGACAGAGCGCTCTCGCCCGCAGCAAGCTTCAGAGGCC 684
Db 1733 AGGACCACTGCGCGTGCAGACAGAGCGCTCTCGCCCGCAGCAAGCTTCAGAGGCC 1674
QY 685 TATGCGGTGAGTGCAGCGGCACAAACCGCTCCCTCAAGGAAGAGGTGTGAGCGGGCCC 744
Db 1673 TATGCGGTGAGTGCAGCGGCACAAACCGCTCCCTCAAGGAAGAGGTGTGAGCGGGCCC 1614
QY 745 GGGAGGAGGAGAGAGCGCAAGGAGGTGACTCGCACTTCCAGTGACACTGATGACA 804
Db 1613 GGGAGGAGGAGAGAGCGCAAGGAGGTGACTCGCACTTCCAGTGACACTGATGACA 1554
QY 805 TTCAGCTCAGATGGAAACAGACCAATGAGCGCAACTCCAACTCGCCCAAGAGAACATGG 864
Db 1553 TTCAGCTCAGATGGAAACAGACCAATGAGCGCAACTCCAACTCGCCCAAGAGAACATGG 1494
QY 865 AGCTGGCTGAGAGCTCAAGAGCTGATTGAGCAGTATGAGCTGGCGAGGAGCATATCG 924
Db 1493 AGCTGGCTGAGAGCTCAAGAGCTGATTGAGCAGTATGAGCTGGCGAGGAGCATATCG 1434
QY 925 ACAAAAGTCTTCAAAACACAAGGACCTTACAAACAGCAGCTGTGATGCTCCAAAGCTCCAGCAGG 984
Db 1433 ACAAAAGTCTTCAAAACACAAGGACCTTACAAACAGCAGCTGTGATGCTCCAAAGCTCCAGCAGG 1374
QY 985 CCCAGGAGATGCTTAAGAGGCGCAGAGAGCGCACCAGCGGGAGAGAGATTTCTCTCTGA 1044
Db 1373 CCCAGGAGATGCTTAAGAGGCGCAGAGAGCGCACCAGCGGGAGAGAGATTTCTCTCTGA 1314
QY 1045 AAGAGCAGTGAAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAAGAGAGCCACCTGA 1104
Db 1313 AAGAGCAGTGAAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAAGAGAGCCACCTGA 1254
QY 1105 AGCAACAGCTTGCCTATATACAGAGAGATTTGAGGAGTTCAGAGACACACTTTCCAAAA 1164
Db 1253 AGCAACAGCTTGCCTATATACAGAGAGATTTGAGGAGTTCAGAGACACACTTTCCAAAA 1194
QY 1165 CGAGCGAGGTATTACACCATTCACAGCAGGAGATGGAAAAAGATGACTAAGAGATCAAGA 1224

1193 GCACGAGGATTACACACATTCAAGCAGGAGATGGAAGATGACTTAAGAGTCAAGA 1134
Db
1225 AGCTGGAGAAAGAAACACACATGATACCGGTCCGGTGGAGAGAGCAACAGCGCCCTGC 1284
Qy
1133 AGCTGGAGAAAGAAACACACATGATACCGGTCCGGTGGAGAGAGCAACAGCGCCCTGC 1074
Db
1285 TTGAGATGGCTGAGGAGAAACACTCCGGGATAAAGAACTGGAGGGCTCCAGGTAAAA 1344
Qy
1073 TTGAGATGGCTGAGGAGAAACACTCCGGGATAAAGAACTGGAGGGCTCCAGGTAAAA 1014
Db
1345 TCCAAACGGCTGGAGAGCTGTCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGA 1404
Qy
1013 TCCAAACGGCTGGAGAGCTGTCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGA 954
Db
1405 GGGTACAGGACCTGAGTGGTGGCCAGAGGCTCCCTCACTGACAGTGGCCCTGAGAGGA 1464
Qy
953 GGGTACAGGACCTGAGTGGTGGCCAGAGGCTCCCTCACTGACAGTGGCCCTGAGAGGA 894
Db
1465 GGGCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGGTCAAGAAAGCGCTTGCT 1524
Qy
893 GGGCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGGTCAAGAAAGCGCTTGCT 834
Db
1525 ACCAGAGAGCAACGAGACAGAGACATCAGCCAGACTGGGGCTCAAGAGCCCACTCCG 1584
Qy
833 ACCAGAGAGCAACGAGACAGAGACATCAGCCAGACTGGGGCTCAAGAGCCCACTCCG 774
Db
1585 CCAGGGCTAGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1644
Qy
773 CCAGGGCTAGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 714
Db
1645 CCTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCATGCTGAGCAGAGATGTTTGA 1704
Qy
713 CCTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCATGCTGAGCAGAGATGTTTGA 655
Db
1705 CCTGGCTGGCATCTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGG 1764
Qy
654 CCTGGCTGGCATCTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGG 595
Db
1765 GCATTTGGCAGGCTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGG 1824
Qy
594 GCATTTGGCAGGCTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGGCACTGG 535
Db
1825 GGGATGGGGTGTACAGATGAAGTCAAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1984
Qy
534 GGGATGGGGTGTACAGATGAAGTCAAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 475
Db
1885 GGGGCTGTCTGTAGTGGCCTCAGTGGCTCAGTGGCTCAGTGGCTCAGTGGCTCAGTGG 1944
Qy
474 GGGGCTGTCTGTAGTGGCCTCAGTGGCTCAGTGGCTCAGTGGCTCAGTGGCTCAGTGG 415
Db
1945 TGCTGTATTGAGGCTCAGAGCCCTCCCTGCCCTT 1979
Qy
414 TGCTGTATTGAGGCTCAGAGCCCTCCCTGCCCTT 380
Db

RESULT 13
HUMHMBGCF/c
LOCUS HUMHMBGCF 1793 bp mRNA linear PRI 25-MAY-1995
DEFINITION Human high molecular weight B cell growth factor mRNA sequence.
ACCESSION L15344
VERSION L15344.1 GI:832913
KEYWORDS B-cell growth factor; interleukin 14.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1793)
AUTHORS Ambrus, J.L. Jr., Pippin, J., Joseph, A., Xu, C., Blumenthal, D.,
Tamayo, A., Claypool, K., McCourt, D., Srikiatchatichorn, A. and
Ford, R. J.
TITLE Identification of a cDNA for a human high-molecular-weight B-cell
growth factor

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (13), 6330-6334 (1993)
MEDLINE 93317675
PUBMED 8327514
REFERENCE 2 (bases 1 to 1793)
AUTHORS Ambrus, J.L. Jr.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-1993) Julian L. Ambrus, Jr., Department of
Medicine, Washington University School of Medicine, St. Louis, MO
63110, USA
COMMENT On May 25, 1995 this sequence version replaced gi:347805.
Original source text: Homo sapiens cDNA to mRNA.
FEATURES
source
1..1793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Namalva"
/tissue_type="Burkitt lymphoma"
ORIGIN
Query Match 24.6%; Score 1156; DB 9; Length 1793;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 703 GGCACAAACGGCTCCCTCAAGGAAGAGTGTGCAGCGGGCCGGGAGGAGGAGGAGC 762
Db 1656 GGCACAAACGGCTCCCTCAAGGAAGAGTGTGCAGCGGGCCGGGAGGAGGAGGAGC 1597
Qy 763 GCAAGGAGTGCACCTCGCACCTTCCAGGTGACACTGAATGACATTCAGCTGCGAGATGGAAC 822
Db 1596 GCNAGGAGTGCACCTCGCACCTTCCAGGTGACACTGAATGACATTCAGCTGCGAGATGGAAC 1537
Qy 823 AGCACAATGAGCGCAACTCCAAAGTTCGCCCAAGAGAACATGGAGCTGGCTGAGAGGTCA 882
Db 1536 AGCACAATGAGCGCAACTCCAAAGTTCGCCCAAGAGAACATGGAGCTGGCTGAGAGGTCA 1477
Qy 883 AGAAGCTGATTCAGCAGTATGAGCTGCGCAGGAGGACATATCGACAAAGTCTTCAACACA 942
Db 1476 AGAAGCTGATTCAGCAGTATGAGCTGCGCAGGAGGACATATCGACAAAGTCTTCAACACA 1417
Qy 943 AGGACCTTCAACAGCAGCTGGTGGATGCCAAGCTCCAGCAGCCCGCAGAGATGTTAAAG 1002
Db 1416 AGGACCTTCAACAGCAGCTGGTGGATGCCAAGCTCCAGCAGCCCGCAGAGATGTTAAAG 1357
Qy 1003 AGGACAGAGCGGCACAGCGGGAGAGGATTTCTCCTGAAAGAGCAGTACAGTCCC 1062
Db 1356 AGGACAGAGCGGCACAGCGGGAGAGGATTTCTCCTGAAAGAGCAGTACAGTCCC 1297
Qy 1063 AGAGGATGTGAGCTGATGAAGCAGCAGAGAGAGCCACCTGAAGCAACAGCTTGCCTAT 1122
Db 1296 AGAGGATGTGAGCTGATGAAGCAGCAGAGAGAGCCACCTGAAGCAACAGCTTGCCTAT 1237
Qy 1123 ACACAGAGAGTTTGAAGGATTCAGAACACACATTTCCAAAAGCAGCGAGTATTCACCA 1182
Db 1236 ACACAGAGAGTTTGAAGGATTCAGAACACACATTTCCAAAAGCAGCGAGTATTCACCA 1177
Qy 1183 CATTCAAGCAGGAGATGAAAAGATGACTAAGAAGATCAAGAAGCTGAGAGAAAGAACCA 1242
Db 1176 CATTCAAGCAGGAGATGAAAAGATGACTAAGAAGATCAAGAAGCTGAGAGAAAGAACCA 1117
Qy 1243 CCATTTACCGGTCCCGTGGGAGAGCAGCAACAGGCCCTCTTGAAGTGGCTGAGAGAGA 1302
Db 1116 CCATTTACCGGTCCCGTGGGAGAGCAGCAACAGGCCCTCTTGAAGTGGCTGAGAGAGA 1057
Qy 1303 AAACAGTCCGGGATAAAGAACTGGAGGGCCCTGCAAGGTAAAAATCAACGGCTGAGAGAGC 1362
Db 1056 AAACAGTCCGGGATAAAGAACTGGAGGGCCCTGCAAGGTAAAAATCAACGGCTGAGAGAGC 997
Qy 1363 TGTCCGGGCACTGCAGACAGAGCGCAATACCTGAACAAGAGGCTACAGACCTGAGTG 1422
Db 996 TGTCCGGGCACTGCAGACAGAGCGCAATACCTGAACAAGAGGCTACAGACCTGAGTG 937
Qy 1423 CTGGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGGCCAGAGGGGCTGGG 1482

[illegible]

1298 GGAGAAAACAGTCCGGGATAAGAACTGGAGGGGCTGCAGGTAAAAATCAAACGGCTGGA 1357
121 GGAGAAAACAGTCCGGGATAAGAACTGGAGGGGCTGCAGGTAAAAATCAAACGGCTGGA 180
1358 GAAGCTGTGCGGGACATGACAGAGAGCGCAATGACCTGAAACAGAGGGTACAGACCT 1417
181 GAAGCTGTGCGGGACATGACAGAGAGCGCAATGACCTGAAACAGAGGGTACAGACCT 240
1418 GAGTGTGTGGCCAGAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGACAGAGGGCC 1477
241 GAGTGTGTGGCCAGAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGACAGAGGGCC 300
1478 TGGGGCTCAAGCACCCAGCTCCCGAGGTCACAGAGCGCTTGTCTACCCAGGACACC 1537
301 TGGGGCTCAAGCACCCAGCTCCCGAGGTCACAGAGCGCTTGTCTACCCAGGACACC 360
1538 GAGCACAGAGCATCAGCCAGACTGGGCTCAAGAGCCCACTCCGCGAGGCGCTAGAG 1597
361 GAGCACAGAGCATCAGCCAGACTGGGCTCAAGAGCCCACTCCGCGAGGCGCTAGAG 420
1598 AGCCTGTGTGGTGCATGTGCGAAGGAGCGGACGCCAGCCAGCGCTGGGCCATAAA 1657
421 AGCCTGTGTGGTGCATGTGCGAAGGAGCGGACGCCAGCGCTGGGCCATAAA 480
1658 AGGCTCCCATGCTGAGCAGGCCATGCTGAAGCCAGGATGTTGACCTGGCTGGCATC 1717
481 AGGCTCCCATGCTGAGCAGGCCATGCTGAAGCCAGGATGTTGACCTGGCTGGCATC 539
1718 TGGCACCTTGCATTTGCAATTTGCAATTTGCGGTGAGTTTACGTACATAGGGCATTTTGAAGG 1777
540 TGGCACCTTGCATTTGCAATTTGCGGTGAGTTTACGTACATAGGGCATTTTGAAGG 599
1778 CTTTGCAAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCAATGGGATGGGGTGT 1837
600 CTTTGCAAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCAATGGGATGGGGTGT 659
1838 GTACAGATGAGTCAGTGGTGTGCTGAGCTGAGCTGAGAGCTTGGAGGGGCTGTATCT 1897
660 GTACAGATGAGTCAGTGGTGTGCTGAGCTGAGAGCTTGGAGGGGCTGTATCT 719
1898 GTAGCTGCATACAGTGAAGTGGCAGAGTGAAGTGTGAGCAATTTCTGCTGCTGATTTGAG 1957
720 GTAGCTGCATACAGTGAAGTGGCAGAGTGAAGTGTGAGCAATTTCTGCTGCTGATTTGAG 779
1958 GCTCAGACCCCTCCCTGCCCTTTACAGAGCTCAAAACAGTAATACACCAAGTCTTGACT 2017
780 GCTCAGACCCCTCCCTGCCCTTTACAGAGCTCAAAACAGTAATACACCAAGTCTTGACT 838
2018 GCATTTGTCTGTGAGCAGGCTTGTGCTGAGTCAAGTCAAGCCCTCCTAGCTGCT-TGGAG 2076
839 GCATTTGTCTGTGAGCAGGCTTGTGCTGAGTCAAGTCAAGCCCTCCTAGCTGCTGAG 898
2077 GCTCCTTTGATTTCTAGACTGGAAGGTTGCTCCTTAGGAGAGCCCTGGCAGGGGCT 2136
899 GCTCCTTTGATTTCTAGACTGGAAGGTTGCTCCTTAGGAGAGCCCTGGCAGGGGCT 958
2137 CAGAGCT-GGATTTTCTGCTGAGCAAGGACCTGGAGAAATGTTTTGCTGGGATGA 2195
959 CAGAGCTGGGATTTCTGCTGAGCAAGGACCTGGAGAAATGTTTTGCTGGGATGA 1018
2196 TGTGCTGGTCAAGAGCCCTTGGGATGCTGCTGCCCTTGGGATGCTGGGATGCAACCA 2255
1019 TGTGCTGGTCAAGAGCCCTTGGGATGCTGCTGCCCTTGGGATGCTGGGATGCAACCA 1078
2256 GGCCAAATGATCTCTCAGTAGCCCTTATCATCAGGTGCTCTCAGGTGCTGCAACCAAT 2315
1079 GGCCAAATGATCTCTCAGTAGCCCTTATCATCAGGTGCTCTCAGGTGCTGCAACCAAT 1138
2316 GATTGACAAGAGATCACCACCAAGGATTTATTTGGAAGGTTTATTTTCTTTATTTCTTTT 2375
1139 GATTGACAAGAGATCACCACCAAGGATTTATTTGGAAGGTTTATTTTCTTTATTTCTTTT 1198

2376 TC--TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCACATCAGTGTGTTATATGAGA 2433
1199 TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCACATCAGTGTGTTATATGAGA 1258
2434 CTTTCCAAAGGAAAGGATGCTGTACAGTGTGCTGGGTGCTTGGCTCCAGTGTGCC 2493
1259 CTTTCCAAAGGAAAGGATGCTGTAGCAGTGTGCTGGGTGCTTGGCTCCAGTGTGCC 1318
2494 ACTCTCTTTCACACACCCCACTTGGCTCTTTTGGCCATCTTGATGCTGAGGTTTCTGTTGG 2553
1319 ACTCTCTTTCACACACCCCACTTGGCTCTTTTGGCCATCTTGATGCTGAGGTTTCTGTTGG 1378
2554 TGAGATCAGGTTGTTTGTGTTAAAGAAAGAAAGGGCTTCTGATGCTTGGCCACAGC 2613
1379 TGAGATCAGGTTGTTTGTGTTAAAGAAAGAAAGGGCTTCTGATGCTTGGCCACAGC 1438
2614 TTACCTGTGGGTTTCAGTCTCTGAGAGGCCACACAGTTCCTCATCAGCACTGTCTCCATG 2673
1439 TTACCTGTGGGTTTCAGTCTCTGAGAGGCCACACAGTTCCTCATCAGCACTGTCTCCATG 1498
2674 CAGCAGTTGTGGGTTCCCATGTCCAGTGTCTCTTGGCTTCATGGGTTTCTGCTTCC 2733
1499 CAGCAGTTGTGGGTTCCCATGTCCAGTGTCTCTTGGCTTCATGGGTTTCTGCTTCC 1558
2734 TSCCTCCACCCCACTCATGTGCAATCTCAAGATTGCTCTGATTTCTATTCTCTGGCACCT 2793
1559 TGCCCCCACCCCACTCATGTGCAATCTCAAGATTGCTCTGATTTCTATTCTCTGGCACCT 1618
2794 CCTCTGCTGTCTTGGGGAATCTACTCTCTCTCTGTGGG-GCCCCATAGCTGTTGCTAA 2852
1619 CCTCTGCTGTCTTGGGGAATCTACTCTCTCTCTGTGGGAGCCCATAGCTGTTGCTAA 1678
2853 CAGGTAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAATCCATAAATGCTGCAG 2912
1679 CAGGTAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAATCCATAAATGCTGCAG 1738
2913 ACAGTTGTTTCTGTGCTGCTTCTACCCCACTCCAGTACATAACTACTACTACTGTGT 2972
1739 ACAGTTGTTTCTGTGCTGCTTCTACCCCACTCCAGTACATAACTACTACTACTGTGT 1798
2973 AGAGCAATCTATATGCTGAATTTCTGCTGTGCAAACTTGCAGGGTATAGCCAGT 3032
1799 AGAGCAATCTATATGCTGAATTTCTGCTGTGCAAACTTGCAGGGTATAGCCAGT 1858
3033 TTTGTGCCAAGCAGTTTTCGGGCAACAGAAATGACTCAGACCAAGATGGAATGATG 3092
1859 TTTGTGCCAAGCAGTTTTCGGGCAACAGAAATGACTCAGACCAAGATGGAATGATG 1918
3093 TAGGCTTTGCTTCTGTGCTGTTTTTTTGA--CTAG--TCATTTCTGTGAGGTCCCTTC 3150
1919 TAGGCTTTGCTTCTGTGCTGTTTTTTTGAAGTAGTTCTATTGCTGTGAGGTCCCTTC 1978
3151 ATCTTCCATACCTAGCCCACTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGTT 3210
1979 ATCTTCCATACCTAGCCCACTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGTT 2038
3211 CACAAAGAAATGTTAAGTACTGAATCATGTGACTGAGACCAAGATGGCAAAATGAATG 3270
2039 CACAAAGAAATGTTAAGTACTGAATCATGTGACTGAGACCAAGATGGCAAAATGAATG 2098
3271 CACCAATTTCTCTCTCTGCCCCCAGGCGAGTACCACCTGATCTGCATCAGAGTTGCC 3330
2099 CACCAATTTCTCTCTCTGCCCCCAGGCGAGTACCACCTGATCTGCATCAGAGTTGCC 2158
3331 TGCTATTCTCTGCTGTATCTTTCATCTAGGTGCCCTCAAGCAGCTGTGTGAGTTGA 3390
2159 TGCTATTCTCTGCTGTATCTTTCATCTAGGTGCCCTCAAGCAGCTGTGTGAGTTGA 2218
3391 GATCTCTGCCATCTCTGCTGAGATGCTGCTGCTGTGAAGTGTTCCTGATGACCTTTT 3450
2219 GATCTCTGCCATCTCTGCTGAGTACTGCTGCTGTGAAGTGTTCCTGATGACCTTTT 2278
3451 TCTTCCCTTTGAATCCCTCT 3471

Db 1199 TTTTTCACATGACAGTGTGTTGTAATTGAGGACCTTCCAAAGGAGGAGGATGCTGTAGC 1258
Qy 2461 AGTGGTGCCTGGGTGCTTGGCTCCAGTGTCCCACTCTCTCTTCAACACCCCACTGGGTCC 2520
Db 1259 AGTGGTGCCTGGGTGCTTGGCTCCAGTGTCCCACTCTCTTCAACACCCCACTGGGTCC 1318
Qy 2521 TTTGCCATCTTGATGCTGAGGTTTCTGTTTGGTGAGATCAGGTGTTGTTGTTGTTGTTAAAGA 2580
Db 1319 TTTGCCATCTTGATGCTGAGGTTTCTGTTTGGTGAGATCAGGTGTTGTTGTTGTTGTTAAAGA 1378
Qy 2581 AAGAAAGGGCTTCGTAGGCTTTGCCAAGCTTACCTGTGGGTTTACGTCTGAGAGG 2640
Db 1379 AAGAAAGGGCTTCGTAGGCTTTGCCAAGCTTACCTGTGGGTTTACGTCTGAGAGG 1438
Qy 2641 CCACACAGTCCCATGACAGTGTCTCCATGACAGAGTGTGGGTCCCATGTCCAGC 2700
Db 1439 CCACACAGTCCCATGACAGTGTCTCCATGACAGAGTGTGGGTCCCATGTCCAGC 1498
Qy 2701 TGCCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCTTCTGCTCCCAACCCCACTGTCAATCCT 2760
Db 1499 TGCCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCTCCCAACCCCACTGTCAATCCT 1558
Qy 2761 CAAGATTGCTGATCTATTTCTGCTGGACCTCCCTGCTGCTGCTTGGGGATCTACCT 2820
Db 1559 CAAGATTGCTGATCTATTTCTGCTGGACCTCCCTGCTGCTGCTTGGGGATCTACCT 1618
Qy 2821 CTTCCTGTGGG-GCCCATAGCTGTGTTCTTAACAGGTAAAGATGAATGAATGAACTATTG 2879
Db 1619 CTTCCTGTGGGAGCCATAGCTGTGTTCTAACAGGTAAAGATGAATGAATGAACTATTG 1678
Qy 2880 ACTGGGCCCCAAGAAATCCATAAATAGCTGCAGACAGTGTGTTCTGCTGCTCTTCTTACC 2939
Db 1679 ACTGGGCCCCAAGAAATCCATAAATAGCTGCAGACAGTGTGTTCTGCTGCTCTTCTTACC 1738
Qy 2940 CCCACTCAGTACATACTATGTAAGTGTAGAGCCATCTATATGCTGAATGTTCT 2999
Db 1739 CCCACTCAGTACATACTATGTAAGTGTAGAGCCATCTATATGCTGAATGTTCT 1798
Qy 3000 GCTGTTGCAAACTTGCCAGGATATTAGCCAGTGTGTTGTGCAAGCAGTGTTCGGGACAA 3059
Db 1799 GCTGTTGCAAACTTGCCAGGATATTAGCCAGTGTGTTGTGCAAGCAGTGTTCGGGACAA 1858
Qy 3060 CAGATGACTCAGACCAAGATGATAGATGGTTAGGGCTTGGCTTCTGCTGTTTCT 3119
Db 1859 CAGATGACTCAGACCAAGATGATAGATGGTTAGGGCTTGGCTTCTGCTGTTTCT 1918
Qy 3120 TTGAA-CTAG-TCAATTGCTGCAAGTCCCTTCACTTCCATACCTAGCCCACTCTTTTA 3177
Db 1919 TTGAAGCTAGTTCATTGTCCTGCAAGTCCCTTCACTTCCATACCTAGCCCACTCTTTTA 1978
Qy 3178 GCCCTTACCTTAATCTCTGATAGTGGTTGATGATGATGATGATGATGATGATGATGATGAT 3237
Db 1979 GCCCTTACCTTAATCTCTGATAGTGGTTGATGATGATGATGATGATGATGATGATGATGAT 2038
Qy 3238 GTGTGACTGAGACCAAGATGGAATGAATGGAACCACTTCTCTCTCTCTCTCTCTCTCTCTCT 3297
Db 2039 GTGTGACTGAGACCAAGATGGAATGAATGGAACCACTTCTCTCTCTCTCTCTCTCTCTCTCT 2098
Qy 3298 GGGCAGGTACCATGATCTGATCAGAGTGGCTGCTGATTTCTCTGGGTATCTCTTCCAT 3357
Db 2099 GGGCAGGTACCATGATCTGATCAGAGTGGCTGCTGATTTCTCTGGGTATCTCTTCCAT 2158
Qy 3358 CTAGGTGCTCAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
Db 2159 CTAGGTGCTCAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218
Qy 3418 TGCTGTCTGTGAAGTGTTCCTGATGACCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3477
Db 2219 TGCTGTCTGTGAAGTGTTCCTGATGACCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2277
Qy 3478 GAGTAGTCTTGTCTTCTTCTGCTCAGTA-GGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3536
Db 2278 GAGTAGTCTTGTCTTCTTCTGCTCAGTA-GGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2336

Qy 3537 CAGGCTAAGCTGGTACAGAGCTCCAA-CTCAGAGCT-TTTCGTAGGCGAGAGAGGTG 3594
Db 2337 CAGGCTAAGCTGGTACAGAGCTCCAACTTCAAGAGTGTTCGTAGGCGAGAGAGGTG 2396
Qy 3595 CAGGAAAGAGGAGGAGGATGATGACCTTCCCTTGAAGAGAGGAAAGGCTTACAGTGG 3654
Db 2397 CAGGAAAGAGGAGGAGGATGATGACCTTCCCTTGAAGAGAGGAAAGGCTTACAGTGG 2456
Qy 3655 CCACATTAATGCTGACTCAGCTTACCTTACCTTAACTGCTTAACTGCTGAGAGGAGTGGAG 3714
Db 2457 CCACATTAATGCTGACTCAGCTTACCTTAACTGCTTAACTGCTGAGAGGAGTGGAG 2516
Qy 3715 CTGCTGATCCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3774
Db 2517 CTGCTGATCCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2576
Qy 3775 TTCCAGGTCAGGAATGTTGGGCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3834
Db 2577 TTCCAGGTCAGGAATGTTGGGCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2636
Qy 3835 TGCTAGACCTTCAGGCCAGTCTGAGATGAGGTTCCCTACCTTTTCTTCTTCTTCTTCTTCTTCT 3894
Db 2637 TGCTAGACCTTCAGGCCAGTCTGAGATGAGGTTCCCTACCTTTTCTTCTTCTTCTTCTTCTTCT 2696
Qy 3895 ACCAAATCAACCAATCACTACAGCTGCTCTCTTCTGCTTTTCCAAAGTAGCCCAAGTCTCT 3954
Db 2697 ACCAAATCAACCAATCACTACAGCTGCTCTCTTCTGCTTTTCCAAAGTAGCCCAAGTCTCT 2756
Qy 3955 GGGCAGATGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4014
Db 2757 GGGCAGATGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2816
Qy 4015 T-GGTCCCACTTGTGACCTTCAGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4073
Db 2817 TGGGTCCTCAGTCTGTA-CTCAGTTTGAAG-CCAGATCTGTTGTTGTTGTTGTTGTTGTTGTT 2874
Qy 4074 TGCTAGCTTTTCTCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4133
Db 2875 TGCTAGCTTTTCTCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2934
Qy 4134 ATAGTAAATTTGTTTTCAGGGCTTGTCCAGCTGAGAGGTTTCTTCCACAGATTTCTGAG 4193
Db 2935 ATAGTAAATTTGTTTTCAGGGCTTGTCCAGCTGAGAGGTTTCTTCCACAGATTTCTGAG 2994
Qy 4194 AGGTGTGAGAGCACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4253
Db 2995 AGGTGTGAGAGCACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3054
Qy 4254 ATGGGCTCAGGCTCAGGCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4313
Db 3055 ATGGGCTCAGGCTCAGGCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3114
Qy 4314 ACCCTCTCTCCACCCCATCAGGTCGCTGAGGCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 4373
Db 3115 ACCCTCTCTCCACCCCATCAGGTCGCTGAGGCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 3174
Qy 4374 TGCTGTCCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4433
Db 3175 TGCTGTCCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3234
Qy 4434 TTTCCGCTTGCACACACAGAGTGGTCACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4493
Db 3235 TTTCCGCTTGCACACACAGAGTGGTCACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3294
Qy 4494 TGCCCTTCTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4553
Db 3295 TGCCCTTCTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3354
Qy 4554 AGCCGGCCACTCAGACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4613
Db 3355 GCGGCGCCACTCAGACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3414

QY 4614 TCACAACAGCCAGGACTGATTTTCATGTATTTTAAACACATTAAATAAGAGTCTGT 4673
 |||||
 Db 3415 TCACAACAGCCAGGACTGATTTTCATGTATTTTAAACACATTAAATAAGAGTCTGT 3474
 |||||
 QY 4674 TGCCTTAAAAAATAAAAAA 4694
 |||||
 Db 3475 TGCCTTAAAAAATAAAAAA 3495
 |||||

RESULT 16
 HSDJ622L5 115756 bp DNA linear PRI 22-NOV-2001
 LOCUS Human DNA sequence from clone RP4-622L5 on chromosome 1p34.2-36.11,
 DEFINITION complete sequence.
 ACCESSION AL049795
 VERSION AL049795.21 GI:17065925
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 115756)
 AUTHORS Frankland, J.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Nov 25, 2001 this sequence version replaced gi:6010175.
 COMMENT During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chrl
 This sequence is the entire insert of clone RP4-622L5 This sequence
 was finished as follows unless otherwise noted: all regions were
 either double-stranded or sequenced with an alternate chemistry or
 covered by high quality data (i.e., phred quality >= 30); an
 attempt was made to resolve all sequencing problems, such as
 compressions and repeats; all regions were covered by at least one
 plasmid subclone or more than one M13 subclone; and the assembly
 was confirmed by restriction digest. RP4-622L5 is from the library
 RPCI-4 constructed by the group of Pieter de Jong. For further
 details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2.

FEATURES
 source
 1..115756
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p34.2-36.11"
 /clone="RP4-622L5"
 /clone_lib="RPCI-4"
 1645..1810
 /note="TIGER1 repeat: matches 2259..2417 of consensus"
 2123..2437
 /note="TIGER1 repeat: matches 1957..2259 of consensus"
 2726..2782
 /note="TIGER1 repeat: matches 1904..1957 of consensus"
 3067..3486
 /note="TIGER1 repeat: matches 1501..1904 of consensus"
 3781..4180
 /note="TIGER1 repeat: matches 1117..1501 of consensus"
 5053..5082
 /note="TIGER1 repeat: matches 1090..1117 of consensus"
 5372..5501
 /note="TIGER1 repeat: matches 966..1090 of consensus"
 5628..5704
 /note="TIGER1 repeat: matches 893..966 of consensus"
 5991..6492
 /note="TIGER1 repeat: matches 411..893 of consensus"
 6800..6835
 /note="TIGER1 repeat: matches 378..411 of consensus"
 6968..7022
 /note="TIGER1 repeat: matches 1..378 of consensus"
 8281..8388
 /note="3 copies 26 mer 79% conserved"
 8301..8360
 /note="30 copies 2 mer ta 81% conserved"
 17674..17777
 /note="4 copies 26 mer 77% conserved"
 17685..17784
 /note="50 copies 2 mer at 82% conserved"
 19385..19542
 /note="match: GSS: Em:AQ306911"
 19551..19574
 /note="12 copies 2 mer tt 100% conserved"
 20865..20922
 /note="29 copies 2 mer at 70% conserved"
 20934..21043
 /note="55 copies 2 mer at 70% conserved"
 20940..21043
 /note="4 copies 26 mer 74% conserved"
 21168..21234
 /note="tRNA-Asp-GAC repeat: matches 1..75 of consensus"
 28263..28691
 /note="match: GSS: Em:AQ317905"
 complement(29964..30447)
 /note="match: GSS: Em:AQ32256"
 30465..30718
 /note="match: STS: Em:G58909
 match: GSS: Em:AQ376964"
 38923..39446
 /note="L1MSc repeat: matches 1546..2095 of consensus"
 39761..39882
 /note="L1MSc repeat: matches 1429..1546 of consensus"
 complement(43254..43726)
 /note="match: GSS: Em:AQ146371"
 45555..46050
 /note="MER44C repeat: matches 1..728 of consensus"
 46055..46468
 /note="match: GSS: Em:AQ040362"
 48914..48982
 /note="MER21B repeat: matches 718..790 of consensus"
 48985..49173
 /note="MER21B repeat: matches 77..263 of consensus"
 49285..49651
 /note="match: GSS: Em:AQ761532"
 52462..53177
 /note="match: GSS: Em:AQ040082"
 54716..54865
 /note="match: GSS: Em:AQ680771"
 60551..60614
 /note="32 copies 2 mer tt 70% conserved"
 complement(62219..62382)
 /note="match: STS: Em:G20585"
 complement(66030..66267)
 /note="match: GSS: Em:AQ0514287"
 complement(66041..66267)
 /note="match: GSS: Em:AQ030288"
 6066..66255
 /note="match: GSS: Em:B69239"
 68882..69541
 /note="CpG island"
 /evidence=not_experimental

```
repeat_region 71427..71448
/note="11 copies 2 mer tt 100% conserved"
misc_feature 73527..74249
/note="match: GSS: Em:AQ892680"
misc_feature 74514..74926
/note="match: GSS: Em:AQ441232"
misc_feature 74532..74981
/note="match: GSS: Em:AQ812740"
misc_feature 74526..75008
/note="match: GSS: Em:AQ091614"
misc_feature complement(74682..75020)
/note="match: STS: Em:L29744"
misc_feature complement(74683..74951)
/note="match: STS: Em:L29729 Em:L29763"
misc_feature 74717..74867
/note="match: STS: Em:G15744"
misc_feature 82692..83089
/note="match: STS: Em:G09447"
misc_feature 83391..83481
/note="match: GSS: Em:AZ310334"
repeat_region 85284..85337
/note="27 copies 2 mer tt 88% conserved"
misc_feature 85339..85718
/note="match: STS: Em:G28385"
misc_feature complement(87250..87612)
/note="match: STS: Em:G06968"
misc_feature complement(87455..87579)
/note="match: STS: Em:G26511"
misc_feature complement(87851..88313)
/note="match: GSS: Em:AQ536603"
repeat_region 89654..89685
/note="16 copies 2 mer tc 87% conserved"
misc_feature 91856..92386
/note="match: GSS: Em:B85187"
repeat_region 92101..92152
/note="26 copies 2 mer tt 71% conserved"
misc_feature 94297..95612
/note="CpG island"
/evidence=not experimental
misc_feature complement(98493..98894)
/note="match: GSS: Em:AQ207463"
misc_feature 103719..104155
/note="match: GSS: Em:AQ543924"
repeat_region 104406..104437
/note="MER57B repeat: matches 303..334 of consensus"
repeat_region 107785..107841
/note="L1MB1 repeat: matches 6111..6171 of consensus"
misc_feature 111249..111767
/note="CpG island"
/evidence=not experimental
misc_feature 113362..113450
/note="match: STS: Em:G16146"
match: GSS: Em:AQ806873"

ORIGIN
Query Match 19.1%; Score 895; DB 9; Length 115756;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2165; Conservative 0; Mismatches 5; Indels 9; Gaps 8;

Qy 1300 AGAAACAGTCCGGGATAAAGAACTGAGGGGCTGCAGGTAAAAATCCAAACGGCTGGAGA 1359
Db 84228 AGAAACAGTCCGGGATAAAGAACTGAGGGGCTGCAGGTAAAAATCCAAACGGCTGGAGA 84287
Qy 1360 AGCTGTGCGGGGACTGACAGACAGAGCGCAATGACCTGAACAAGAGGGTACAGCACTGA 1419
Db 84288 AGCTGTGCGGGGACTGACAGACAGAGCGCAATGACCTGAACAAGAGGGTACAGCACTGA 84347
Qy 1420 GTGCTGTGGCGAGGGTCCCTCACTACAGTGGCCCTGAGAGAGGCGGAGGGGCGCTG 1479
Db 84348 GTGCTGTGGCGAGGGTCCCTCACTACAGTGGCCCTGAGAGAGGCGGAGGGGCGCTG 84407
Qy 1480 GGGCTCAAGCACCCAGCTCCCGCCAGGGTCAACAAGCGCCCTTGTACCCAGGACACCGA 1539
```

```
Db 84408 GGGCTCAAGCACCCAGCTCCCGCCAGGGTCAACAAGCGCCCTTGTACCCAGGACACCGA 84467
Qy 1540 GCACAGAGCATCAGCCAGACTGGGCTCAAGAGCCCACTCCGCGAGGGCTTAGAGAG 1599
Db 84468 GCACAGAGCATCAGCCAGACTGGGCTCAAGAGCCCACTCCGCGAGGGCTTAGAGAG 84527
Qy 1600 CCTGGTGTGGGTCACTGTGGGAAGGGAGCGGAGCCAGCCAGCCAGCCAGCCATAAAG 1659
Db 84528 CCTGGTGTGGGTCACTGTGGGAAGGGAGCGGAGCCAGCCAGCCAGCCATAAAG 84587
Qy 1660 GCTCCCATGCTGAGCAGCCCATTCCTGAAGCAAGATGTTTTCACCTGGCTGGCATCTG 1719
Db 84588 GCTCCCATGCTGAGCAGCCCATTCCTGAAGCAAGATGTTTTCACCTGGCTGGCATCTG 84646
Qy 1720 SCACCTTCAATTTTGGATTTTGGGTTTGGGTTTGGGTTTACGTACATAGGGATTTTGAAGGCC 1779
Db 84647 GCACCTTCAATTTTGGATTTTGGGTTTGGGTTTGGGTTTACGTACATAGGGATTTTGAAGGCC 84706
Qy 1780 TTGCAAAATGCATTTATACCTGTAAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTGT 1839
Db 84707 TTGCAAAATGCATTTATACCTGTAAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTGT 84766
Qy 1840 ACAGATGAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGCTGTCTATCTGT 1899
Db 84767 ACAGATGAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGCTGTCTATCTGT 84826
Qy 1900 AGCTGCCATCACAGTGAAGTGGCAGAAGTGAAGTCTGAGCATTTCTCTGTCTCATTTGAGGC 1959
Db 84827 AGCTGCCATCACAGTGAAGTGGCAGAAGTGAAGTCTGAGCATTTCTCTGTCTCATTTGAGGC 84886
Qy 1960 TCAGACCCCTCCCTGCTTTCAGAGCTCAAAAACAAGTAATACCAAGTCTTGACTGC 2019
Db 84887 TCAGACCCCTCCCTGCTTTCAGAGCTCAAAAACAAGTAATACCAAGTCTTGACTGC 84945
Qy 2020 ATTGCTCTTGAGCAGAGGCTTGTCTGTGAGCTCAAGCCCTCTAGCTGTCTTGGAGGC 2078
Db 84946 ATTGCTCTTGAGCAGAGGCTTGTCTGTGAGCTCAAGCCCTCTAGCTGTCTTGGAGGC 85005
Qy 2079 TCCTTTGATTTCTTAGACCTGAAAAGGTGTCCTTAGCAGAGCCCTGGCAGGCGCTCA 2138
Db 85006 TCCTTTGATTTCTTAGACCTGAAAAGGTGTCCTTAGCAGAGCCCTGGCAGGCGCTCA 85065
Qy 2139 GAGCT-GGATTTCTGCTGGAACAGAGGACCTGGAGAAATGTTTTCGCTGGGATGATG 2197
Db 85066 GAGCTGGGGATTTCTGCTGGAACAGAGGACCTGGAGAAATGTTTTCGCTGGGATGATG 85125
Qy 2198 TGCTGCTCAGAGAGCCCTTGGGCTATCGCTTCCCTGCTGCTTGGTGTAGTGCAGGACGAGG 2257
Db 85126 TGCTGCTCAGAGAGCCCTTGGGCTATCGCTTCCCTGCTGCTTGGTGTAGTGCAGGACGAGG 85185
Qy 2258 CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGGCTGCAAAATGA 2317
Db 85186 CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGGCTGCAAAATGA 85245
Qy 2318 TTGACAGAGATCAACCAAGGATTAATTTCTGAAGGTGTTTTTTTCTTTATTTCTTTTC 2377
Db 85246 TTGACAGAGATCAACCAAGGATTAATTTCTGAAGGTGTTTTTTTCTTTATTTCTTTTC 85305
Qy 2378 --TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTGTAGTGTCCAGTGTCCAC 2435
Db 85306 TTTT-----TTTTTTTTTTTTTTTTTTTTTTTGACATGACAGTGTGTTTGTATTGAGGACC 85365
Qy 2436 TTCCAGGAAAGGATGCTGTACAGTGTGCTGGGTGCTGCTGCTCCAGTGTCCAC 2495
Db 85366 TTCCAGGAAAGGATGCTGTAGCAGTGTGCTGGGTGCTGCTGCTCCAGTGTCCAC 85425
Qy 2496 CTCCTTACCCACCTTGGCTCTTTTGCATCTTGATGTGAGGTTCCTGTTGGTG 2555
Db 85426 CTCCTTACCCACCTTGGCTCTTTTGCATCTTGATGTGAGGTTCCTGTTGGTG 85485
Qy 2556 AGATCAGGTTCTTGTGTTAAAGAAAGAGGAGGCTTCTGATGGCTTGGCCACCAAGCTT 2615
Db 85486 AGATCAGGTTCTTGTGTTAAAGAAAGAGGAGGCTTCTGATGGCTTGGCCACCAAGCTT 85545
```

2616 ACCTGTGGGTTTCAGTCTCCTGAGAGGCGCACACAGTTCCCATCAGCACTGTCTCCATGCA 2675
Db 85946 ACCTGTGGGTTTCAGTCTCCTGAGAGGCGCACACAGTTCCCATCAGCACTGTCTCCATGCA 85605
Qy 2676 GCAGTTCTGGGTTCCCATGAGTCCAGTCCCTCTTTGGCTTCATGGGTTTTCCTCTCCG 2735
Db 85606 GCAGTTCTGGGTTCCCATGAGTCCAGTCCCTCTTTGGCTTCATGGGTTTTCCTCTCCG 85665
Qy 2736 CCCCCACCCACATGTCGAATCTCAAGATTTGTCTGATTTCTATTTCTGCGACCTCC 2795
Db 85666 CCCCCACCCACATGTCGAATCTCAAGATTTGTCTGATTTCTATTTCTGCGACCTCC 85725
Qy 2796 CTGCTGTCTCTGGGATTTCTATTTCTCTGTGTGGG-GCCCATAGCTGTTTCTTAACA 2854
Db 85726 CTGCTGTCTCTGGGATTTCTATTTCTCTGTGTGGGAGCCCATAGCTGTTTCTTAACA 85785
Qy 2855 GGTAAAGAAATGAATGAATGAATGACTGGGCCCCAGAAATCCATAAATGGCTGCAGAC 2914
Db 85786 GGTAAAGAAATGAATGAATGAATGACTGGGCCCCAGAAATCCATAAATGGCTGCAGAC 85845
Qy 2915 AGTTGTTTCTGTGCTCTGTTCTACCCCACTCCAGTACATAACTACTATGTACTGTGTAG 2974
Db 85905 AGTTGTTTCTGTGCTCTGTTCTACCCCACTCCAGTACATAACTACTATGTACTGTGTAG 85905
Qy 2975 AGCATTTCTATGCTGAATGTTCTGCTGTGCAACTTCCAGGTTATTTAGCAGTGT 3034
Db 85906 AGCATTTCTATGCTGAATGTTCTGCTGTGCAACTTCCAGGTTATTTAGCAGTGT 85965
Qy 3035 TGTGCCAAGCAGTTTTCGGGCAACAAGATGACTCAGACCAAGATGGATAGGATGTTA 3094
Db 85966 TGTGCCAAGCAGTTTTCGGGCAACAAGATGACTCAGACCAAGATGGATAGGATGTTA 86025
Qy 3095 GGGCTTTGCTTCTGCTGTTTCTTTGAA-CTAG-TCATTTGCTGCGAGGTCCTTCAT 3152
Db 86026 GGGCTTTGCTTCTGCTGTTTCTTTGAAGCTAGTTTCATTTGCTGCGAGTCCCTTCAT 86085
Qy 3153 CTTCATACCTAGCCACTCTTTTAGCCCTTACCTTAATCTCTCAGATAAGTTGTTCA 3212
Db 86086 CTTCATACCTAGCCACTCTTTTAGCCCTTACCTTAATCTCTCAGATAAGTTGTTCA 86145
Qy 3213 CAAGAATGTTAAGTACTGAATCATGTGCTGAGACCAAGATGGCAATGAATGGCA 3272
Db 86146 CAAGAATGTTAAGTACTGAATCATGTGCTGAGACCAAGATGGCAATGAATGGCA 86205
Qy 3273 CACCATTTCTCTCTCTGCGCCAGGAGTACCACTGATCTGCATCAGAGTTGCTTG 3332
Db 86206 CACCATTTCTCTCTCTGCGCCAGGAGTACCACTGATCTGCATCAGAGTTGCTTG 86265
Qy 3333 CTATTTCTGCTGATCTCTTCACTCTAGGTGCGCTCAAGCAGCTGTGTGAGTTGAGA 3392
Db 86266 CTATTTCTGCTGATCTCTTCACTCTAGGTGCGCTCAAGCAGCTGTGTGAGTTGAGA 86325
Qy 3393 TCTCTGCACTCTGGCTGAGATCTGCTGCTGTGAAAGTGTTCCTGACCTTTTTC 3452
Db 86326 TCTCTGCACTCTGGCTGAGATCTGCTGCTGTGAAAGTGTTCCTGACCTTTTTC 86385
Qy 3453 TTCCCTTTTGAATCCCTCT 3471
Db 86386 TTCCCTTTTGAATCCCTCT 86404

RESULT 17
AR409342
LOCUS AR409342 22255 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 51 from patent US 6632923.
ACCESSION AR409342
VERSION AR409342.1 GI:40160130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22255)
Unclassified

AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 51 14-OCT-2003;
FEATURES Location/Qualifiers
source 1..22255
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 18.0%; Score 844; DB 6; Length 22255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 6; Indels 9; Gaps 8;
Qy 1300 AGAAAACAGTCGGGATTAAGAACTGGAGGGCTGACAGTAAATCCAAAGCTGGAGA 1359
Db 16637 AGAAAACAGTCGGGATTAAGAACTGGAGGGCTGACAGTAAATCCAAAGCTGGAGA 16696
Qy 1360 AGCTGTGCGGGCCTGACAGACAGAGCGCAATGACCTGAAAGAGGTACAGACCTGA 1419
Db 16697 AGCTGTGCGGGCCTGACAGACAGAGCGCAATGACCTGAAAGAGGTACAGACCTGA 16756
Qy 1420 GTGCTGTGGGCGAGGGCTCCCTCACTGACAGTGGGCCCTGAGAGGAGGCGGCGCTG 1479
Db 16757 GTGCTGTGGGCGAGGGCTCCCTCACTGACAGTGGGCCCTGAGAGGAGGCGGCGCTG 16816
Qy 1480 GGGCTCAAGCACCCAGCTCCCGCAGGGTCAAGAGCGCTTGTACCCAGGAGCACGA 1539
Db 16817 GGGCTCAAGCACCCAGCTCCCGCAGGGTCAAGAGCGCTTGTACCCAGGAGCACGA 16876
Qy 1540 GCACAGAGCATCAGGCGCAGACTGGGCCCTCAAGAGCCACCTCCGCCAGGGCTAGAG 1599
Db 16877 GCACAGAGCATCAGGCGCAGACTGGGCCCTCAAGAGCCACCTCCGCCAGGGCTAGAG 16936
Qy 1600 CTTGTTTGGGTCACTGTGGAGAGGAGCGGACCCAGAGCGCTTGCCCAATAAAG 1659
Db 16937 CTTGTTTGGGTCACTGTGGAGAGGAGCGGACCCAGAGCGCTTGCCCAATAAAG 16996
Qy 1660 GCTCCCATGCTCAGCAGCCCATTTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTG 1719
Db 16997 GCTCCCATGCTCAGCAGCCCATTTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTG 17055
Qy 1720 GCATTTGCAATTTTGGATTTTGTGGTTCAGTTTACGTACATAGGCAATTTTGAAGGCC 1779
Db 17056 GCATTTGCAATTTTGGATTTTGTGGTTCAGTTTACGTACATAGGCAATTTTGAAGGCC 17115
Qy 1780 TTGCAATGCAATTTTACCTGTAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTG 1839
Db 17116 TTGCAATGCAATTTTACCTGTAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTG 17175
Qy 1840 ACAGATGAAGTCAGTGGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCTCTGT 1899
Db 17176 ACAGATGAAGTCAGTGGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCTCTGT 17235
Qy 1900 AGCTGCAATCAAGTGTGTCAGAGTGTGCTGAGCATTTCTGCTGTGATTTGAGGC 1959
Db 17236 AGCTGCAATCAAGTGTGTCAGAGTGTGCTGAGCATTTCTGCTGTGATTTGAGGC 17295
Qy 1960 TCAGACCCCTCCCTCCCTTTCAGAGCTCAAAACAAAGTAATACACAAAGTCTTCACTGC 2019
Db 17296 TCAGACCCCTCCCTCCCTTTCAGAGCTCAAAACAAAGTAATACACAAAGTCTTCACTGC 17354
Qy 2020 ATTGTCTTGTGAGCAGGGCTTGTGTCAGCTCAGGCGCTTCTAGTGTCT-TGAGGC 2078
Db 17355 ATTGTCTTGTGAGCAGGGCTTGTGTCAGCTCAGGCGCTTCTAGTGTCTTCTGAGGC 17414
Qy 2079 TCTTTGATTTCTAGACTGGAAGGTGTCCTTAGCAGAGCCCTGCGAGGGCGCTCA 2138
Db 17415 TCTTTGATTTCTAGACTGGAAGGTGTCCTTAGCAGAGCCCTGCGAGGGCGCTCA 17474
Qy 2139 GAGCT-TGGATTTCTGCTGCAACAAGAGGAGCTTGAGAGATGTTTTTTCGTGGGATGATG 2197
Db 17475 GAGCTGGGATTTGCTGCTGCAACAAGAGGAGCTTGAGAGATGTTTTTTCGTGGGATGATG 17534

QY 2198 TGCTGGTCAGAGCCCTTGGGCACTGCTCCCTGCCCTTGGTGTAGTGCAGGACGAG 2257
Db 17535 TGCTGGTCAGAGCCCTTGGGCACTGCTCCCTGCCCTTGGTGTAGTGCAGGACGAG 17594
QY 2258 CCAATGATGCTTCTCAGTAGGCTTATCATTCACAGGTGCTCTCTAGCCCTGCACAAATGA 2317
Db 17595 CCAATGATGCTTCTCAGTAGGCTTATCATTCACAGGTGCTCTCTAGCCCTGCACAAATGA 17654
QY 2318 TTGACAGAGATCACCCAAAGGATATTTCTGAAGGTGTTTTCTTTATTTCTTTTTC 2377
Db 17655 TTGACAGAGATCACCCAAAGGATATTTCTGAAGGTGTTTTCTTTATTTCTTTTTC 17714
QY 2378 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACATGACAGGTGTTGATTGAGGACC 2435
Db 17715 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACATGACAGGTGTTGATTGAGGACC 17774
QY 2436 TTCCAAGAAAGGATGCTGTACAGTGTGCTGGGTGCTGGCTCCAGTGTGCCAC 2495
Db 17775 TTCCAAGAAAGGATGCTGTACAGTGTGCTGGGTGCTGGCTCCAGTGTGCCAC 17834
QY 2496 CTCCTTACCAACCCACTTGGCTCTTGGCATCTTGCATGCTGAGGTTTCTGTTGGTG 2555
Db 17835 CTCCTTACCAACCCACTTGGCTCTTGGCATCTTGCATGCTGAGGTTTCTGTTGGTG 17894
QY 2556 AGATCAGGTTGTTGTGTAAGAAAGAAAGGCTTCTGATGGCTTTTGCACAAAGCTT 2615
Db 17895 AGATCAGGTTGTTGTGTAAGAAAGAAAGGCTTCTGATGGCTTTTGCACAAAGCTT 17954
QY 2616 ACCTGTGGGTTTCACTGCTGAGAGGCCACACAGTTCCCATCAGACTGCTCCATGCA 2675
Db 17955 ACCTGTGGGTTTCACTGCTGAGAGGCCACACAGTTCCCATCAGACTGCTCCATGCA 18014
QY 2676 GCAGTGTGCTGGTCCCATGCTCCAGTGCCTTTTGGCTTTCATGGGTTTTCTGCTCCTG 2735
Db 18015 GCAGTGTGCTGGTCCCATGCTCCAGTGCCTTTTGGCTTTCATGGGTTTTCTGCTCCTG 18074
QY 2736 CCCCCACCCACACATGTCGAATCTCAGATTTGCTCTGATTTCTATTTCTGCGACCTCC 2795
Db 18075 CCCCCACCCACACATGTCGAATCTCAGATTTGCTCTGATTTCTATTTCTGCGACCTCC 18134
QY 2796 CTGCTGTCTTGGGATTTCTACTTCTTCTGTGTGGG-GCCCATAGCTGTTCTCTAAACA 2854
Db 18135 CTGCTGTCTTGGGATTTCTACTTCTTCTGTGTGGGAGCCCATAGCTGTTCTCTAAACA 18194
QY 2855 GGTAAAGAAATGAATTAACATTTGATGTGGGCCCCAGAAATCCATAAAATGGCTGCAGAC 2914
Db 18195 GGTAAAGAAATGAATTAACATTTGATGTGGGCCCCAGAAATCCATAAAATGGCTGCAGAC 18254
QY 2915 AGTTGTTTCTGTGCTCTTACCCCACTCCAGTACATAAATCTATCTATCTGTGTAG 2974
Db 18255 AGTTGTTTCTGTGCTCTTACCCCACTCCAGTACATAAATCTATCTATCTGTGTAG 18314
QY 2975 AGCCATTCTATATGCTGAATGTTTCTGTGTTGCAAACTTGCCAGGGTATTAGCCAGTGT 3034
Db 18315 AGCCATTCTATATGCTGAATGTTTCTGTGTTGCAAACTTGCCAGGGTATTAGCCAGTGT 18374
QY 3035 TGTGCCAAGCAGTTTTCGGGGAACAAGATGACTCAGACAAGATGGATAGATGGTTA 3094
Db 18375 TGTGCCAAGCAGTTTTCGGGGAACAAGATGACTCAGACAAGATGGATAGATGGTTA 18434
QY 3095 GGGCTTTGCTTCTGCTGTTTTCTTTTGA-CTAG-TCAATGTCTCCAGGTCCCTTTCAT 3152
Db 18435 GGGCTTTGCTTCTGCTGTTTTCTTTGAGCTAGTTCAATGTCTCCAGGTCCCTTTCAT 18494
QY 3153 CTTCACATCTAGCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA 3212
Db 18495 CTTCACATCTAGCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA 18554
QY 3213 CAAGAATGTTAAGTACTGTAATCATGTGCTAGACACAGATGCGCAATGAATGGCA 3272
Db 18555 CAAGAATGTTAAGTACTGTAATCATGTGCTAGACACAGATGCGCAATGAATGGCA 18614

QY 3273 CACCAATTTCTCTCTCTGCCCCAGGCGAGTACCACTGATCTGTCATCAGATTCCTG 3332
Db 18615 CACCAATTTCTCTCTCTGCCCCAGGCGAGTACCACTGATCTGTCATCAGATTCCTG 18674
QY 3333 CTATTTCTGTGTATCTTTCATCTAGTGTGCTTCAAGCAGCTGTGTAGTGTGAGA 3392
Db 18675 CTATTTCTGTGTATCTTTCATCTAGTGTGCTTCAAGCAGCTGTGTAGTGTGAGA 18734
QY 3393 TCTCTGCCATCTCTGCTGAGATCTGCTGCTGTAAGTGTTCCTCATGACCTTTTC 3452
Db 18735 TCTCTGCCATCTCTGCTGAGATCTGCTGCTGTAAGTGTTCCTCATGACCTTTTC 18794
QY 3453 TTCCCTTTGAATCCCTCT 3471
Db 18795 TTCCCTTTGAATCCCTCT 18813
RESULT 18
AX239608
LOCUS 22255 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 51 from Patent WO0164874.
ACCESSION AX239608
VERSION AX239608.1 GI:15797283
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 51 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
Location/Qualifiers
source 1..22255
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 18.0%; Score 844; DB 6; Length 22255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 6; Indels 9; Gaps 8;
QY 1300 AGAAACAGTCCGGATAAAGAACTGAGGGCTGAGGCTAAATAATCCACGGCTGAGA 1359
Db 16637 AGAAACAGTCCGGATAAAGAACTGAGGGCTGAGGCTAAATAATCCACGGCTGAGA 16596
QY 1360 AGCTGTGCGGGCACTGTCAGACAGAGCGCAATGACCTGAACAAGAGGTACAGGACCTGA 1419
Db 16697 AGCTGTGCGGGCACTGTCAGACAGAGCGCAATGACCTGAACAAGAGGTACAGGACCTGA 16756
QY 1420 GTGTGTGGCCAGGGCTCCCTCACTGACAGTGCCTTGAAGAGGCGCAGAGGGGCTG 1479
Db 16757 GTGTGTGGCCAGGGCTCCCTCACTGACAGTGCCTTGAAGAGGCGCAGAGGGGCTG 16816
QY 1480 GGGCTCAGACACCCAGCTCCCGGGTCAAGAGCGCTTGTACCCAGGACACCGA 1539
Db 16817 GGGCTCAGACACCCAGCTCCCGGGTCAAGAGCGCTTGTACCCAGGACACCGA 16876
QY 1540 GCACAGAAGCATCAGGCCAGACTGGGCTCAAGAGCCACCTCCCGCAGGCTTAGAGAG 1599
Db 16877 GCACAGAAGCATCAGGCCAGACTGGGCTCAAGAGCCACCTCCCGCAGGCTTAGAGAG 16936
QY 1600 CTGTGTGTTGGTCAATGCTGGAGAGGAGCGGACCGCCAGCGCTTGGCCCAATAAG 1659
Db 16937 CTGTGTGTTGGTCAATGCTGGAGAGGAGCGGACCGCCAGCGCTTGGCCCAATAAG 16996
QY 1660 GCTCCCATGCTGAGCAGGCCCATTTGCTGAAGCCAGGATGTTTCTTGACCTGGCTGGCATCTG 1719
Db 16997 GCTCCCATGCTGAGCAGGCCCATTTGCTGAAGCCAGGATGTTTCTTGACCTGGCTGGCATCTG 17055

QY 1720 GCATTCGAAATTTGGATTTTGGGTCAGTTTATACATAGGGGCAATTTGCAAGGCC 1779
Db 17056 GCATTCGAAATTTGGATTTTGGGTCAGTTTATACATAGGGGCAATTTGCAAGGCC 17115
QY 1780 TTGCAATGATTTATACCTGTAAGTGTACAGTGGGCTTCATTTGGGGATGGGGGTGTGT 1839
Db 17116 TTGCAATGATTTATACCTGTAAGTGTACAGTGGGCTTCATTTGGGGATGGGGGTGTGT 17175
QY 1840 ACAGATGAAGTCAGTGGCTTTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGT 1899
Db 17176 ACAGATGAAGTCAGTGGCTTTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGT 17235
QY 1900 AGCTGCCATCACAGTGAAGTGGCAGAGTCACTTGAGCAATTTCTGTCTGATTTGAGGC 1959
Db 17236 AGCTGCCATCACAGTGAAGTGGCAGAGTCACTTGAGCAATTTCTGTCTGATTTGAGGC 17295
QY 1960 TCAGACCCCTCCCTGCCCTTTTCAGAGCTCAAAACAGTAATACACAAAGTCTTGACTGC 2019
Db 17296 TCAGACCCCTCCCTGCCCTTTTCAGAGCTCAAAACAGTAATACACAAAGTCTTGACTGC 17354
QY 2020 ATTTGCTGTGAGCAGGGCTTGTGTCAGTCAAGTCAAGCCCTCTAGCTGCT-TGAGGC 2078
Db 17355 ATTTGCTGTGAGCAGGGCTTGTGTCAGTCAAGTCAAGCCCTCTAGCTGCTCTGAGGC 17414
QY 2079 TCCTTTGATTTCTAGACCTTGGAAAGTGTCCCTAGCAGAGCCCTGGCAGGCGCTCA 2138
Db 17415 TCCTTTGATTTCTAGACCTTGGAAAGTGTCCCTAGCAGAGCCCTGGCAGGCGCTCA 17474
QY 2139 GAGCT-GGGATTTCTGCTGCTGGAAACAGGACCTGGAGATTTTGGCTGGGATGATG 2197
Db 17475 GAGCTGGGATTTCTGCTGCTGGAAACAGGACCTGGAGATTTTGGCTGGGATGATG 17534
QY 2198 TGCTGTCAGAGGCCCTTGGGCACTGCTCCCTGCGCTTGGTAGTCCAGGACGAG 2257
Db 17535 TGCTGTCAGAGGCCCTTGGGCACTGCTCCCTGCGCTTGGTAGTCCAGGACGAG 17594
QY 2258 CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGCCTGCAAAATGA 2317
Db 17595 CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGCCTGCAAAATGA 17654
QY 2318 TTGCAAGAGATACCCAAAGGATTTATCTGAAAGTGTGTTTTTCTTTATTTCTTTTC 2377
Db 17655 TTGCAAGAGATACCCAAAGGATTTATCTGAAAGTGTGTTTTTCTTTATTTCTTTTC 17714
QY 2378 -TTTTTTTTTTTTTTTTTTTTTTTTTTTGGCAGATGAGTGTGTTGATGAGGAC 2435
Db 17715 TTTTTTTTTTTTTTTTTTTTTTTTTTTTGGCAGATGAGTGTGTTGATGAGGAC 17774
QY 2436 TTCCAAGAAAGGATGCTGTACAGTGTGCTGGGTGCTGGCTCCAGTGTCCAC 2495
Db 17775 TTCCAAGAAAGGATGCTGTAGCAGTGTGCTGGGTGCTGGCTCCAGTGTCCAC 17834
QY 2496 CTCCTTCCACACCCCTGCTGCTCTTTGCACTGCTGATGCTGAGGTTTCTGTTGGTG 2555
Db 17835 CTCCTTCCACACCCCTGCTGCTCTTTGCACTGCTGATGCTGAGGTTTCTGTTGGTG 17894
QY 2556 AGATCAGGTTGTTGTGTTAAAGAAAGGAGGCTTCTGATGGCTTTGCCCAAGCTT 2615
Db 17895 AGATCAGGTTGTTGTGTTAAAGAAAGGAGGCTTCTGATGGCTTTGCCCAAGCTT 17954
QY 2616 ACCTGTGGTTTGTGCTGAGAGGCCACACAGTTCCCATGAGCTGCTCCATGCA 2675
Db 17955 ACCTGTGGTTTGTGCTGAGAGGCCACACAGTTCCCATGAGCTGCTCCATGCA 18014
QY 2676 GCAGTTGCTGGTCCCATGCTCAGCTGCTCTTTGGCTTTCATGGGTTTTCTGCTTCTG 2735
Db 18015 GCAGTTGCTGGTCCCATGCTCAGCTGCTCTTTGGCTTTCATGGGTTTTCTGCTTCTG 18074
QY 2736 CCCCCACCCACATGTCGAATCTCAAGATTTGCTGATTTCTATTTCTGTCGACCTCC 2795
Db 18075 CCCCCACCCACATGTCGAATCTCAAGATTTGCTGATTTCTATTTCTGTCGACCTCC 18134
QY 2796 CTGCTGTCTGGGATTTCTACTTCTCTGCTGGG-GCCCATGCTGTTGCTCAACA 2854

Db 18135 CTGCTGTCTTGGGGATTTCTACTTCTCTCTGTGGAGCCCATAGCTGTGTGCTAACA 18194
QY 2855 GGTAAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAAATCCATAAAATGGCTCGACAG 2914
Db 18195 GGTAAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAAATCCATAAAATGGCTCGACAG 18254
QY 2915 AGTTGTTTCTGCTGCTGCTTCTACCCCACTCCAGTCAATAACTACTATGTAAGTGTAG 2974
Db 18255 AGTTGTTTCTGCTGCTGCTTCTACCCCACTCCAGTCAATAACTACTATGTAAGTGTAG 18314
QY 2975 AGCCATTTCTATGCTGAATTTCTGCTGTTGCAAACTTGCAGGGTATTAGCCAGTGT 3034
Db 18315 AGCCATTTCTATGCTGAATTTCTGCTGTTGCAAACTTGCAGGGTATTAGCCAGTGT 18374
QY 3035 TGTGCAAGCAGTTTTCGGGGACAAGAAATGACTCAGACCAAGATGGATAGGATGTTA 3094
Db 18375 TGTGCAAGCAGTTTTCGGGGACAAGAAATGACTCAGACCAAGATGGATAGGATGTTA 18434
QY 3095 GGGCTTTCCTGCTGCTGTTTTCTTTGAA-CTAG-TCAATGTCCTGAGGTCCCTTCAT 3152
Db 18435 GGGCTTTCCTGCTGCTGTTTTCTTTGAACTAGTCTTCAATGCTCTGAGGTCCCTTCAT 18494
QY 3153 CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTGTGTTCA 3212
Db 18495 CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTGTGTTCA 18554
QY 3213 CAAAGAAATTTAAGTACTGAATCATGTGCTGAGCAGAGATGGCAATGAATGCA 3272
Db 18555 CAAAGAAATTTAAGTACTGAATCATGTGCTGAGCAGAGATGGCAATGAATGCA 18614
QY 3273 CACCATTTCTCTCTCTGCCCCCAGGAGTACCACTGATCTGCATCAGAGTTGCCGTG 3332
Db 18615 CACCATTTCTCTCTCTGCCCCCAGGAGTACCACTGATCTGCATCAGAGTTGCCGTG 18674
QY 3333 CTATTCTCTGCTGATTCCTTCACTAGTGGCTCCCTCAAGCAGTGTGTGAGTGTGAGA 3392
Db 18675 CTATTCTCTGCTGATTCCTTCACTAGTGGCTCCCTCAAGCAGTGTGTGAGTGTGAGA 18734
QY 3393 TCTCTGCCATCTCTGCTGAGATGCTGCTGCTGAGTGTTCCTGAGTGTTCCTGAGTGTTC 3452
Db 18735 TCTCTGCCATCTCTGCTGAGATGCTGCTGCTGAGTGTTCCTGAGTGTTCCTGAGTGTTC 18794
QY 3453 TTCCCTTTTGAATCCCTCT 3471
Db 18795 TTCCCTTTTGAATCCCTCT 18813

RESULT 19
HSM802997

LOCUS 2713 bp mRNA linear PRI 20-MAR-2002
DEFINITION Homo sapiens mRNA; cDNA DKPZp667F1912 (from clone DKPZp667F1912).

ACCESSION AL713662

VERSION AL713662.1 GI:19584349

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 2713)

Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.

Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152

Martinsried, GERMANY

Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKPZ); Email: s.wiemann@dkfz-heidelberg.de;

sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German

Genome Project.

This clone (DKPZp667F1912) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further

information about the clone and the sequencing project is available

Db	1513	TACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAGTTGGTTGCACAAAGA	1572
Qy	3219	ATGTTAAGTACTGAATCATCTGTGACTGACAGACAGAGATGGCAATGAATGGCACACCAT	3278
Db	1573	ATGTTAAGTACTGAATCATCTGTGACTGACAGACAGAGATGGCAATGAATGGCACACCAT	1632
Qy	3279	TTCTCCCTTCCTCCGCCCCCGGAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATTTC	3338
Db	1633	TTCTCCCTTCCTCCGCCCCCGGAGGTACCACTGATCTGCATCAGAGTTGCCTGCTATTTC	1692
Qy	3339	TCTGGTGTATCTTCAATCTAGTCCCTCCCTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTG	3398
Db	1693	TCTGGTGTATCTTCAATCTAGTCCCTCCCTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTG	1752
Qy	3399	CCATCTCTGCTGAGTACTGCTGCTGCTGAGAGTTTCCCATGACCTTTTCTTCCCTCC	3458
Db	1753	CCATCTCTGCTGAGTACTGCTGCTGCTGAGAGTTTCCCATGACCTTTTCTTCCCTCC	1812
Qy	3459	TTTGAATCCCTCT 3471	
Db	1813	TTTGAATCCCTCT 1825	
RESULT 21			
AF070546		1477 bp	linear
LOCUS	AF070546	1477 bp	linear
DEFINITION	Homo sapiens clone 24607 mRNA sequence.		
ACCESSION	AF070546		
VERSION	AF070546.1	GI:3387908	
KEYWORDS	FLI_CDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Anderson, B., Wentland, M.A., Ricafrente, J.Y., Liu, W. and Gibbs, R.A.		
TITLE	A 'double adaptor' method for improved shotgun library construction		
JOURNAL	Anal. Biochem. 236 (1), 107-113 (1996)		
MEDLINE	96207227		
PUBMED	8619474		
REFERENCE	2 (bases 1 to 1477)		
AUTHORS	Yu, W., Anderson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W., Ricafrente, J.Y., Wentland, M.A., Lennon, G. and Gibbs, R.A.		
TITLE	Large-scale concatenation cDNA sequencing		
JOURNAL	Genome Res. 7 (4), 353-358 (1997)		
MEDLINE	97264341		
PUBMED	9110174		
REFERENCE	3 (bases 1 to 1477)		
AUTHORS	Yu, W. and Gibbs, R.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-JUN-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA		
FEATURES	Location/Qualifiers		
source	1. 1477		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="I.M.A.G.E. Consortium clone ID 24607"		
	/sex="female"		
	/tissue_type="brain"		
	/clone_lib="INIB"		
	/dev_stage="infant"		
ORIGIN			
Query Match	11.5%;	Score 538;	DB 9; Length 1477;
Best Local Similarity	99.5%;	Pred. No. 4.3e-295;	
Matches 1118;	Conservative 0;	Mismatches 2;	Indels 4; Gaps 4;
Qy	3575	TTTCTAGGCGAGAGAGGTGCAGGAGAGCGAGAGATGACCTTCCCTTGAAGAG	3634
Db	347	TTTCTAGGCGAGAGAGGTGCAGGAGAGCGAGAGATGACCTTCCCTTGAAGAG	406
Qy	3635	AGGGGAAGGCGCTACAGTGGCCACATAATTTGCTGACTCACATCTCAGCTACCTCTTAA	3694


```
VERSION AX126027.1 GI:34532358
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagaatsuma, M., Murakawa, K., Kanehori, K.,
Takashi-Fujii, A., Oshina, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 4054)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..4054
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RAB-4028904"
/tissue="testis"
/clone_lib="TEST14"
/note="Cloning vector: pME18SFL3"
ORIGIN
Query Match 7.4%; Score 346; DB 9; Length 4054;
Best Local Similarity 99.6%; Pred. No. 3.6e-185;
Matches 446; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4233 TTTCATGAGTTATCGACCATGGCTGAGCTCAGCAGGCTTTCTGTAGGAGACTGTAT 4292
Db 4052 TTTCATGAGTTATCGACCATGGCTGAGCTCAGCAGGCTTTCTGTAGGAGACTGTAT 3993
Qy 4293 TTCTGTAAGATGTTATTTAAACCTCTCTCCACCCATCAGGTGGCCCTGAGGGCTGAC 4352
Db 3992 TTCTGTAAGATGTTATTTAAACCTCTCTCCACCCATCAGGTGGCCCTGAGGGCTGAC 3933
Qy 4353 CCGAGGCGCAGTGCAGCTGCTGTGTCCAGGGGGAGGCGCTGCTGAGCTGAT 4412
Db 3932 CCGAGGCGCAGTGCAGCTGCTGTGTCCAGGGGGAGGCGCTGCTGAGCTGAT 3873
Qy 4413 TCTCAGCTGCTGCCCGCCAGCCCTTTCCGCTTTCACACAGAGGTGGTCAACCCAGGGA 4472
Db 3872 TCTCAGCTGCTGCCCGCCAGCCCTTTCCGCTTTCACACAGAGGTGGTCAACCCAGGGA 3813
Qy 4473 CAGCAGGCGCAGTGCAGCTGCTGTGTCCAGGGGGAGGAGCTGCTGTCCCTGT 4532
Db 3812 CAGCAGGCGCAGTGCAGCTGCTGTGTCCAGGGGGAGGAGCTGCTGTCCCTGT 3753
Qy 4533 AACTGCTTCTTATGGCCCAACCCGGCCACTCAGACTTGTTCAGCTGCACCTGGCAGC 4592
Db 3752 AACTGCTTCTTATGGCCCAACCCGGCCACTCAGACTTGTTCAGCTGCACCTGGCAGC 3693
Qy 4593 TTTTGTGCTCCTTTGGTATTCACACAGCCAGGACTGATTTGATGATTTTAAAC 4652
Db 3692 TTTTGTGCTCCTTTGGTATTCACACAGCCAGGACTGATTTGATGATTTTAAAC 3633
```

```
QY 4653 CACATTAATAAAGAGTCTGTGCTTAA 4680
Db 3632 CACATTAATAAAGAGTCTGTGCTTAA 3605

RESULT 24
AX893880
LOCUS AX893880
DEFINITION Sequence 9743 from Patent EP1033401.
ACCESSION AX893880
VERSION AX893880.1 GI:40048764
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
Expressed sequence tags and encoded human proteins
Patent: EP 1033401-A 9743 06-SEP-2000;
Genset (FR)
FEATURES
source
1..316
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 5.1%; Score 239; DB 6; Length 316;
Best Local Similarity 100.0%; Pred. No. 5.5e-124;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2144 GGGATTTCCTGCTGGACAAGGACCTGGAGAAATGTTTTCGGTGGGATGATGTGCTGG 2203
Db 78 GGGATTTCCTGCTGGACAAGGACCTGGAGAAATGTTTTCGGTGGGATGATGTGCTGG 137
QY 2204 TCAGGAGCCCTTGGGCAATCGTCCCTGCTGCTTGGTAGTCAGGACGAGGCAATG 2263
Db 138 TCAGGAGCCCTTGGGCAATCGTCCCTGCTGCTTGGTAGTCAGGACGAGGCAATG 197
QY 2264 ATGCTTTCAGTAGCTTATTCATTCACAGGTGCTCTAGCTGCACAAATGATTGACA 2323
Db 198 ATGCTTTCAGTAGCTTATTCATTCACAGGTGCTCTAGCTGCACAAATGATTGACA 257
QY 2324 AGAGATCAACCCAAAGGATTTCTGAAGGTGTTTTTCTTATTTCTTTCTTTT 2382
Db 258 AGAGATCAACCCAAAGGATTTCTGAAGGTGTTTTTCTTATTTCTTTCTTTT 316

RESULT 25
BD029413
LOCUS BD029413
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD029413
VERSION BD029413.1 GI:22571155
KEYWORDS JP 2001269182-A/5659.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 316)
Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
Sequence tag and encoded human protein
Patent: JP 2001269182-A 5659 02-OCT-2001;
GENSET
REFERENCE OS Homo sapiens (human)
AUTHORS PN JP 2001269182-A/5659
TITLE PD 02-OCT-2001
JOURNAL DP 24-FEB-2000 JP 2000118773
COMMENT PR 26-FEB-1999 US 60/422487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
```

C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC

FEATURES FH Key Location/Qualifiers.

LOCUS
1.316

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 5.1%; Score 239; DB 6; Length 316;

Best Local Similarity 100.0%; Pred. No. 5.5e-124; Indels 0; Gaps 0;

Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2144 GGGATTTCTGCTGGGAAACAGGACCTGGAGAAATGTTTTGCTGGGATGATGCTGG 2203

Db 78 GGGATTTCTGCTGGGAAACAGGACCTGGAGAAATGTTTTGCTGGGATGATGCTGG 137

QY 2204 TCAGAGGCCCTTGGGATCGCTTCCCTGCGCTTGGTAGTGCAGGACGAGCCAAATG 2263

Db 138 TCAGAGGCCCTTGGGATCGCTTCCCTGCGCTTGGTAGTGCAGGACGAGCCAAATG 197

QY 2264 ATGCTTCTCAGTAGCCTTATCATTCACAGGTGCTCTCTAGCCTGCAACAATGATTGACA 2323

Db 198 ATGCTTCTCAGTAGCCTTATCATTCACAGGTGCTCTCTAGCCTGCAACAATGATTGACA 257

QY 2324 AGAGATCACCAAGGATATTTCTGAAGGTGTTTTTCTTTATTTCTTTTCTTTT 2382

Db 258 AGAGATCACCAAGGATATTTCTGAAGGTGTTTTTCTTTATTTCTTTTCTTTT 316

RESULT 26

AX396823

LOCUS

DEFINITION

Sequence 1038 from Patent WO0212328.

AX396823

ACCESSION

AX396823.1

VERSION

1

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

King, G.E., Meagher, M.J., Xu, J. and Secrist, H.

TITLE

Compositions and methods for the therapy and diagnosis of colon

cancer

JOURNAL

Patent: WO 0212328-A 1038 14-FEB-2002;

CORIXA CORPORATION (US)

FEATURES

source

1.235

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches 235; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY 4164 CTGAGAGCTCATGCTCCACAGATTTCTGAGAGTGTGTCAGGACACTTTTTTTTATTT 4223

Db 1 CTGAGAGCTCATGCTCCACAGATTTCTGAGAGTGTGTCAGGACACTTTTTTTTATTT 60

QY 4224 GTTGTGTTTGTTCATGAGGTATTTCGACCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 4283

Db 61 GTTGTGTTTGTTCATGAGGTATTTCGACCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 120

QY 4284 GACTGTTATTCTTAAGATGTTTATTAACTCTCCACCCCATCAGGTGGCCCTG 4343

Db 121 GACTGTTATTCTTAAGATGTTTATTAACTCTCCACCCCATCAGGTGGCCCTG 180

QY 4344 AGGGCTACCCCGAGGACCCAGTGGAGCTGCTGGTGTCTCCACGGGGAGGCCCAAGG 4398

Db 181 AGGGCTACCCCGAGGACCCAGTGGAGCTGCTGGTGTCTCCACGGGGAGGCCCAAGG 235

RESULT 27

AX834462

LOCUS

DEFINITION

Sequence 1586 from Patent EP1347046.

AX834462

ACCESSION

AX834462.1

VERSION

1

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Isoigai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,

Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,

Tamachika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and

Masuo, Y.

TITLE

Full-length cDNA sequences

Patent: EP 1347046-A 1586 24-SEP-2003;

Research Association for Biotechnology (JP)

FEATURES

source

1.2194

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 4.9%; Score 230; DB 6; Length 2194;

Best Local Similarity 99.6%; Pred. No. 8.7e-119;

Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAAAGCAGCCCGAGGACCAACCGGAGCAGGACCCGAGGAGCCCGAGGAGCGGCCAGCC 60

Db 112 CAAAAGCAGCCCGAGGACCAACCGGAGCAGGACCCGAGGAGCGGCCAGCC 171

QY 61 AGCGGGCTCTCTGAGTAGAAGCAAGGTCCTCCGCGAGCAGGCTCTCTCGAAGCGG 120

Db 172 AGCGGGCTCTCTGAGTAGAAGCAAGGTCCTCCGCGAGCAGGCTCTCTCGAAGCGG 231

QY 121 AGCGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTCAGGAGTGA 180

Db 232 AGCGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTCAGGAGTGA 291

QY 181 GCGGCCAACTGGAGACATCTGAGCATACTGTGTGCAATAACCCAGGGGGCCCG 240

Db 292 GCGGCCAACTGGAGACATCTGAGCATACTGTGTGCAATAACCCAGGGGGCCCG 351

QY 241 GCGAGGATGGGCGCACAGGGTGAGCGGCTGAACCCGAGAT 281

Db 352 GCGAGGATGGGCGCACAGGGTGAGCGGCTGAACCCGAGAT 392

RESULT 28

AK096939

LOCUS

DEFINITION

Homo sapiens cDNA FLJ39620 fis, clone SMINT200114.

AK096939

ACCESSION

AK096939.1

VERSION

1

KEYWORDS

oligo capping; fis (full insert sequence).

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,

Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Watanabe, M.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,

Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,

Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,

LOCUS	AR425015	421 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 16512 from patent US 6639063.				
ACCESSION	AR425015				
VERSION	AR425015.1	GI:40180125			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 421)				
AUTHORS	Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.				
TITLE	EST's and encoded human proteins				
JOURNAL	Patent: US 6639063-A 16512 28-OCT-2003;				
FEATURES	Location/Qualifiers				
source	1..421				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	3.7%;	Score 173;	DB 6;	Length 421;	
Best Local Similarity	99.7%;	Pred.No.3.3e-86;			
Matches 293;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;	
QY	4217 TTTATTGTGTTGTTGTTTCCATGAGGTTATCGGACCATGGGCTGAGCTCAGGCACCTTC	4276			
Db	1 TTTATTGTGTTGTTGTTTCCATGAGGTTATCGGACCATGGGCTGAGCTCAGGCACCTTC	60			
QY	4277 TGTAGGAGACTGTTATTCTGTAAAGATGGTTATTTAAACCTCTCCACCCATCAGGT	4336			
Db	61 TGTAGGAGACTGTTATTCTGTAAAGATGGTTATTTAAACCTCTCCACCCATCAGGT	120			
QY	4337 GGCCTCTGAGGCTGACCCGAGGCCAGTGGCTGCTGTGTCACGGGGAGGGCCAA	4396			
Db	121 GGCCTCTGAGGCTGACCCGGA-GCCAGTGGAGTGCTGTGTCACGGGGAGGGCCAA	179			
QY	4397 GGCCTCTGAGCTGATTCACAGCTGCTGCCCAGGCTTTCGCCTTGCACGACAGAG	4456			
Db	180 GGCCTCTGAGCTGATTCACAGCTGCTGCCCAGGCTTTCGCCTTGCACGACAGAG	239			
QY	4457 GTGGTCACCCCGGGACAGCAGGACCTGCTCTCTTGGCCCTTCTTGGGGAA	4510			
Db	240 GTGGTCACCCCGGGACAGCAGGACCTGCTCTTGGCCCTTCTTGGGGAA	293			
RESULT 32					
BD120568		421 bp	DNA	linear	PAT 18-SBP-2002
LOCUS	BD120568				
DEFINITION	EST and encoded human protein.				
ACCESSION	BD120568				
VERSION	BD120568.1	GI:23215478			
KEYWORDS	JP 2002010789-A/12645.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 421)				
AUTHORS	Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.				
TITLE	EST and encoded human protein				
JOURNAL	Patent: JP 2002010789-A 12645 15-JAN-2002;				
COMMENT	GENSET CORP				
	OS Homo sapiens (human)				
	PN JP 2002010789-A/12645				
	PD 15-JAN-2002				
	PF 07-AUG-2000 JP 2000280989				
	PR 05-AUG-1999 US 60/147499				
	PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE				PI
	GIORDANO				
	PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC				
	C12N1/21,				
	PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC				
	C12N15/00				
	CC EST and encoded human protein				
	FT Key Location/Qualifiers				
	FT source 1..421				

FEATURES	source	Location/Qualifiers	/organism='Homo sapiens' (human)'
ORIGIN			
Query Match		3.7%;	Score 173; DB 6; Length 421;
Best Local Similarity		99.7%;	Pred. No. 3.3e-86;
Matches 293;	Conservative	0;	Mismatches 0; Indels 1; Gaps 1;
QY	4217	TTTATTGTGTGTTGTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACCTTTC	4276
Db	1	TTTATTGTGTGTTGTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACCTTTC	60
QY	4277	TGTAGGAGACTGTTATTTCTGTAAGATGGTTATTAAACCTCTCTCCACCCCATCACGGT	4336
Db	61	TGTAGGAGACTGTTATTTCTGTAAGATGGTTATTAAACCTCTCTCCACCCCATCACGGT	120
QY	4337	GGCCCTGAGGCTGACCCGGAGGCCAGTGAGCTGCTGTGTTCACGGGGAGGGCCAA	4396
Db	121	GGCCCTGAGGCTGACCCGGAGGCCAGTGAGCTGCTGTGTTCACGGGGAGGGCCAA	179
QY	4397	GGCCTGCTGAGCTGATTCTCCAGCTGCTGCCCCAGCCCTTCCGCCCTTGACAGCACAGAG	4456
Db	180	GGCCTGCTGAGCTGATTCTCCAGCTGCTGCCCCAGCCCTTCCGCCCTTGACAGCACAGAG	239
QY	4457	GTGGTCACCCAGGGACAGCCAGGACCTGTCTCTTGGCCCTTCTGGGGGAA	4510
Db	240	GTGGTCACCCAGGGACAGCCAGGACCTGTCTCTTGGCCCTTCTGGGGGAA	293
RESULT 33			
LOCUS	AX331983	441 bp	DNA linear PAT 09-JAN-2002
DEFINITION	Sequence 2492 from Patent WO0194629.		
ACCESSION	AX331983		
VERSION	AX331983.1	GI:18122617	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Horrikan, S., Soppet D.R. and Weaver, Z.		
TITLE	Cancer gene determination and therapeutic screening using signature		
JOURNAL	Patent: WO 0194629-A 2492 13-DEC-2001;		
FEATURES	source	Location/Qualifiers	
ORIGIN			
Query Match		3.5%;	Score 166; DB 6; Length 441;
Best Local Similarity		100.0%;	Pred. No. 3.4e-82;
Matches 166;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	3306	ACCACTGATCTGCATCAGAGTTCCTGCTATCTCTGGTGATCTCTTCCATCTAGGTGC	3365
Db	43	ACCACTGATCTGCATCAGAGTTCCTGCTATCTCTGGTGATCTCTTCCATCTAGGTGC	102
QY	3366	CCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGCTGTCC	3425
Db	103	CCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGCTGTCC	162
QY	3426	TGTGAAGTGTTCATGACCTTTTCTCCCTTTTGAATCCCTCT	3471
Db	163	TGTGAAGTGTTCATGACCTTTTCTCCCTTTTGAATCCCTCT	208


```
QY 4641 TGTATTTTAAACCACTTAATAAAGAGTCTGTGTCCTT 4679
Db 39 TGTATTTTAAACCACTTAATAAAGAGTCTGTGTCCTT 1
RESULT 36
LOCUS AR199533 4722 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 14 from patent US 6355451.
ACCESSION AR199533
VERSION AR199533.1 GI:20249607
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4722)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6355451-A 14 12-MAR-2002;
FEATURES
source
Location/Qualifiers
1..4722
/mol_type="genomic DNA"
ORIGIN
Query Match 1.3%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 6.8e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1514
RESULT 37
LOCUS AR374685 4722 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 14 from patent US 6605588.
ACCESSION AR374685
VERSION AR374685.1 GI:40077500
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4722)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6605588-A 14 12-AUG-2003;
FEATURES
source
Location/Qualifiers
1..4722
/mol_type="genomic DNA"
ORIGIN
Query Match 1.3%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 6.8e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1514
RESULT 38
LOCUS AR409320 4722 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 14 from patent US 6632923.
ACCESSION AR409320
VERSION AR409320.1 GI:40160108
KEYWORDS
```

```
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4722)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 14 14-OCT-2003;
FEATURES
source
Location/Qualifiers
1..4722
/mol_type="genomic DNA"
ORIGIN
Query Match 1.3%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 6.8e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1514
RESULT 39
LOCUS AX239571 4722 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 14 from Patent WO0164874.
ACCESSION AX239571
VERSION AX239571.1 GI:15797253
KEYWORDS
SOURCE
ORGANISM
Oryctolagus cuniculus (rabbit)
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 14 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
source
Location/Qualifiers
1..4722
/mol_type="Oryctolagus cuniculus"
/mol_type="unassigned DNA"
/db_xref="taxon:9986"
61..1734
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC88294.1"
/db_xref="GI:15797254"
/db_xref="REMTREMBL:CAC88294"
/translation="MKNDKNGAKQNPFSQPGPAGAGAGCGRPPAPAREAE
GASSQAPRPGCAQTAQPGALCDVSELSRQLEDILSTYCVNNQAGPQVQGE
PPEDAEKRAYVARNGEPFGTPVNGEKETSKAEPTGTEIRTSDEVGDRHRRPQ
EKKAKGLGKEITLLMOTLNTLSTPEKLAALCKKYALBLEHNSOKMLLKQKOS
OLVQEKDHLRGHESKAILARSKLRLQHNRSLEKGVQVQAREREEKREKVTSH
QMTLNDIQLQMEQHNERNRSKLQENNELAELKLIQVELREHIDIKVFKHKLQO
QLVDAGQQAQEMLKAEERHQREKDFLLKEAVESQRMCHLMKQOETHLKOQLALYE
KPEFQNTLSKSEVFTTFKQEMKMTKKIKLEKETTMYKRWESSNKALLEABEK
TLRDELGLQVKIQRLKQLQTERNDLNKRVQDLISAGQGVSDSGSPERRPEPA
TTSKEQGVGPGAQVPSPRATDASCCAGAPSTEASGQTGPQEPPTATA"
ORIGIN
Query Match 1.3%; Score 59; DB 6; Length 4722;
Best Local Similarity 100.0%; Pred. No. 6.8e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGTT 1514
RESULT 40
```


BD056447 4722 bp DNA linear PAT 27-AUG-2002
 LOCUS Novel low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis,
 BD056447
 ACCESSION
 VERSION GI:22602053
 KEYWORDS JP 2001506983-A/5.
 SOURCE Aequorea victoria
 ORGANISM Aequorea victoria
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
 Aequoreidae; Aequorea.
 1 (bases 1 to 4722)
 Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 Novel low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 Patent: JP 2001506983-A 5 29-MAY-2001;
 BOSTON HEART FOUNDATION INC
 PN JP 2001506983-A/5
 PD 29-MAY-2001
 PF 26-NOV-1997 JP 1998524870
 PR 27-NOV-1996 US 60/031930, 03-JUN-1997 US 60/048547 PI
 ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC
 A61K38/04, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC
 C07H21/00,
 PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES Location/Qualifiers
 source
 1. .4722
 /organism="Aequorea victoria"
 /mol_type="genomic DNA"
 /db_xref="taxon:6100"
 ORIGIN
 Query Match 1.3%; Score 59; DB 6; Length 4722;
 Best Local Similarity 100.0%; Pred. No. 6.8e-21;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1350 CGCGTGAGAGCTGTCCGGGCACTGCACACAGAGCGCAATGACCTGACACAGAGGT 1408
 Db 1456 CGCGTGAGAGCTGTCCGGGCACTGCACACAGAGCGCAATGACCTGACACAGAGGT 1514
 RESULT 41
 G26511/c 142 bp DNA linear STS 02-JUN-1996
 LOCUS human STS TIGR-A004Z24, sequence tagged site.
 DEFINITION
 G26511
 VERSION GI:1348743
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 142)
 Hudson,T.
 Whitehead Institute/MIT Center for Genome Research; Physically
 Mapped STSs
 Unpublished (1995).
 JOURNAL
 COMMENT
 Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu
 Primer A: AATCATCAAAATCAAGTCC
 Primer B: CTTCTGCTGCTGTAAGT
 STS size: 142
 PCR Profile:

Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng
 Primer: each 5 pM
 dNTPs: each 4 mM
 Tag Polymerase: 0.025 units/ul
 Total Vol: 20 ul
 Buffer:
 MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCL: 10 mM
 pH: 9.3
 Derived from dbEST (genbank accession R73477).
 Location/Qualifiers
 1. .142
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="160.0 CR from top of Chr9 linkage group"
 STS
 primer_bind 1. .142
 primer_bind complement(125. .142)
 ORIGIN
 Query Match 1.0%; Score 49; DB 11; Length 142;
 Best Local Similarity 100.0%; Pred. No. 3e-15;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4580 CTGCACCTGCGACGCTTTTGTCTCTCTTTGGGTATTCACACAGCCAGG 4628
 Db 68 CTGCACCTGCGACGCTTTTGTCTCTCTTTGGGTATTCACACAGCCAGG 20
 RESULT 42
 AC111774
 LOCUS Rattus norvegicus clone CH230-171G2, WORKING DRAFT SEQUENCE, 10
 DEFINITION
 AC111774 245095 bp DNA linear HTG 09-NOV-2002
 Rattus norvegicus clone CH230-171G2, WORKING DRAFT SEQUENCE, 10
 unordered pieces.
 AC111774
 AC111774.4 GI:24818709
 VERSION
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 245095)
 Muzny,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
 Allen,C., Allien,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpthy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kovar, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresuhewa, L., Lousleged, H., Lorzado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, I.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwunonu, G., Olarnpunagcon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., P.L.L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, W.A., Reigh, R., Reilly, B., Reilly, M., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vallas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 245095)
Worley, K.C.

Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 245095)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23321701.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOBQ
Center clone name: CH230-171G2
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 153773 bases at least Q40
Consensus quality: 160170 bases at least Q30
Consensus quality: 164659 bases at least Q20
Estimated insert size: 160182; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 69762: contig of 69762 bp in length
69763 69862: gap of unknown length
69863 73689: contig of 3827 bp in length
73690 73789: gap of unknown length
73790 88447: contig of 14658 bp in length
88448 88547: gap of unknown length
88548 124613: contig of 36066 bp in length
124614 124713: gap of unknown length
124714 179270: contig of 54557 bp in length
179271 179370: gap of unknown length
179371 224496: contig of 45126 bp in length
224497 224596: gap of unknown length
224597 228136: contig of 3540 bp in length
228137 228336: gap of unknown length
228337 239465: contig of 11229 bp in length
239466 239566: gap of unknown length
239567 243842: contig of 4277 bp in length
243843 243942: gap of unknown length
243943 245095: contig of 1153 bp in length.

FEATURES
Location/Qualifiers
1..245095
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-171G2"
misc_feature 88548..89924
/note="wgs_contig"
misc_feature 91432..93819
/note="wgs_contig"
misc_feature 116848..118080
/note="wgs_contig"
misc_feature 124714..126005
/note="wgs_contig"
misc_feature 228337..229783
/note="wgs_contig"
misc_feature 239566..240945
/note="wgs_contig"

ORIGIN
Query Match 1.0%; Score 49; DB 2; Length 245095;
Best Local Similarity 100.0%; Pred.No. 4.5e-15; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0

QY 1371 GCATGTCAGACAGAGCGCAATGACCTGAACAGAGGTTACAGGACCTGA 1419
|||||
DB 145976 GCATGTCAGACAGAGCGCAATGACCTGAACAGAGGTTACAGGACCTGA 146024
|||||

RESULT 43
BC060227
LOCUS BC060227 4702 bp mRNA linear ROD 02-DEC-2003
DEFINITION Mus musculus cDNA clone MGC:77972 IMAGE:30357541, complete cds.
ACCESSION BC060227
VERSION BC060227.1 GI:38614270
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4702)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blackley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnurch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4702)

Strausberg, R.

Direct Submission

Submitted (20-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: casabs@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

CDNA Library Preparation: M. Bento Soares, University of Iowa

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu

Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

FEATURES

Source

1. 4702

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="MGC:77972 IMAGE:30357541"

/tissue_type="Brain"

/clone_lib="NIH_BMAP_GH0"

/lab_host="DH10B"

/note="Vector: pYX-ASC"

220..1884

CDS

/codon_start=1

/product="Unknown (protein for MGC:77972)"

/protein_id="AAH60227.1"

/db_xref="GI:38614271"

/translation="MNQDKNGPKAKSHSKGSPGQREAGPEGAGHGRPQTAPGAAE
GTSOAPGKTEGARAAAPGALCVSELSROLEDILSTYCVDNNGGPAEAGQAE
PTPEDEKSTVAANGPEPGIPVNGEKETSKEGPETEIRASDEVDGDRHRRPO
EKKAKGLKEITLLQTLTSTPEKLAALCKKYAELLEHRNSOKMKLLQKQS
QLVQDKDHGESHKAVLAKSKLESCLRELQNRHSKEQGVQVAREEKEKRETVSH
FOYINDILOQEHNRNSKURQENLEAKRLKLEIQEYELREEDIKVFHKHQLQ
QLVADKLQAQAEERHQRKEFLKAEVSRQMCLEMKQETHLQKQLALYTE
KFEFQNTLSKSEVFTFKQENKMTKKIKLEKETMYRWRSSNKKALLEAEKK
TVRDLKGLQVQKIORLEKLCRALQTERNDLNKRVDLTAGGITDIGSRREPEATTAS
KEQGVSPGQAPASSPRATDAPCCSAPSTGTAGTGPGEPIFATA"

ORIGIN

Query Match 1.0%; Score 47; DB 10; Length 4702;

Best Local Similarity 100.0%; Pred. No. 5.le-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 780 CACTTCCAGGTGACACTGAATGACATTCAGTCAGTCAGTGGACAGCA 826
DB 1045 CACTTCCAGGTGACACTGAATGACATTCAGTCAGTGGACAGCA 1091

RESULT 44

AL671759/c
LOCUS AL671759.10 141790 bp DNA linear ROD 12-JUL-2002
DEFINITION Mouse DNA sequence from clone RP23-151A15 on chromosome 4, complete sequence.

ACCESSION AL671759

KEYWORDS HTG. GI:22138725

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 141790)

REFERENCE

CORBY, N.
Direct Submission
Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Aug 8, 2002 this sequence version replaced gi:21732087.

COMMENT

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-151A15 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

Source

1. 141790

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="4"

/clone="RP23-151A15"

/clone_lib="RPI-23"

ORIGIN

Query Match 1.0%; Score 47; DB 10; Length 141790;
Best Local Similarity 100.0%; Pred. No. 6.le-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 780 CACTTCCAGGTGACACTGAATGACATTCAGTCAGTGGACAGCA 826
DB 62204 CACTTCCAGGTGACACTGAATGACATTCAGTCAGTGGACAGCA 62158

COMMENT

On Aug 19, 2000 this sequence version replaced gi:8117638.

```
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-697E18
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 171244 bases at least Q40
Consensus quality: 173243 bases at least Q30
Consensus quality: 174393 bases at least Q20
Insert size: 175126; sum-of-contigs
Quality coverage: 10.66x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1. 44460 contig of 44460 bp in length
44561 62100 contig of 17540 bp in length
62201 80631 contig of 18431 bp in length
80732 98278 contig of 17547 bp in length
98379 110236 contig of 11858 bp in length
110337 121052 contig of 10716 bp in length
121153 130490 contig of 9338 bp in length
130591 139290 contig of 8700 bp in length
139391 145488 contig of 6098 bp in length
145589 154222 contig of 8634 bp in length
154323 161014 contig of 6692 bp in length
161115 167741 contig of 6627 bp in length
167742 167841 contig of 100 bp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 44460: contig of 44460 bp in length
* 44561 44560: gap of 100 bp
* 44561 62100: contig of 17540 bp in length
* 62101 62200: gap of 100 bp
* 62201 80631: contig of 18431 bp in length
* 80632 80731: gap of 100 bp
* 80732 98278: contig of 17547 bp in length
* 98279 98378: gap of 100 bp
* 98379 110236: contig of 11858 bp in length
* 110337 110336: gap of 100 bp
* 110337 121052: contig of 10716 bp in length
* 121053 121152: gap of 100 bp
* 121153 130490: contig of 9338 bp in length
* 130491 130590: gap of 100 bp
* 130591 139290: contig of 8700 bp in length
* 139291 139390: gap of 100 bp
* 139391 145488: contig of 6098 bp in length
* 145489 145588: gap of 100 bp
* 145589 154222: contig of 8634 bp in length
* 154223 154322: gap of 100 bp
* 154323 161014: contig of 6692 bp in length
* 161015 161114: gap of 100 bp
* 161115 167741: contig of 6627 bp in length
* 167742 167841: gap of 100 bp
-----
RESULT 45
AP000665/c 103829 bp DNA linear PRI 15-MAR-2003
LOCUS Homo sapiens genomic DNA, chromosome 11q clone: CMB9-46G18, complete
DEFINITION sequences.
ACCESSION AP000665
VERSION AP000665.5 GI:13094209
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens genomic DNA
Published Only in Database (1999)
2 (bases 1 to 103829)
Direct Submission
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Submitted (04-NOV-1999) Masahira Hattori, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Feb 21, 2001 this sequence version replaced gi:12381912.
FEATURES
Location/Qualifiers
source
1. 103829
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clones="CMB9-46G18"
ORIGIN
Query Match 0.9%; Score 44; DB 9; Length 103829;
Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 4289 TTATTTCTGAAGATGGTATTATTAACCTCCCTCCACCCCATCA 4332
|||||
Db 8497 TTATTTCTGAAGATGGTATTATTAACCTCCCTCCACCCCATCA 84454
|||||
RESULT 46
AP001192/c 176526 bp DNA linear HTG 19-AUG-2000
LOCUS Homo sapiens chromosome 11 clone RP11-697E18 map 11q23, WORKING
DEFINITION DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION AP001192
VERSION AP001192.3 GI:9857605
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 176526)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 176,526 genomic DNA of 11q23
Published Only in DataBase (2000)
2 (bases 1 to 176526)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Submitted (18-FEB-2000) Masahira Hattori, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-42-778-9923, Fax: 81-42-778-9924)
```

* 167842 172932: contig of 4951 bp in length
 * 172933 172932: gap of 100 bp
 * 175180 175180: contig of 2288 bp in length
 * 175181 175280: gap of 100 bp
 * 175281 176526: contig of 1246 bp in length.

FEATURES

source
 1. 176526
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11Q23"
 /clone="RP11-697E18"
 1. 44460
 /note="assembly_fragment"
 44561..62100
 /note="assembly_fragment clone_end:77 vector_side:right"
 62201..80631
 /note="assembly_fragment"
 80732..98278
 /note="assembly_fragment"
 98379..110236
 /note="assembly_fragment"
 110337..121052
 /note="assembly_fragment"
 121153..130490
 /note="assembly_fragment"
 130591..139290
 /note="assembly_fragment"
 139391..145488
 /note="assembly_fragment clone_end:SP6 vector_side:right"
 145589..154222
 /note="assembly_fragment"
 154323..161014
 /note="assembly_fragment"
 161115..167741
 /note="assembly_fragment"
 167842..172792
 /note="assembly_fragment"
 172893..175180
 /note="assembly_fragment"
 175281..176526
 /note="assembly_fragment"
 ORIGIN
 Query Match 0.9%; Score 44; DB 2; Length 176526;
 Best Local Similarity 100.0%; Pred. No. 3.2e-12;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4289 TTATTTCTGTAAGATGGTATTATACCTTCCTCCACCCCATCA 4332
 |||||
 Db 5885 TTATTTCTGTAAGATGGTATTATACCTTCCTCCACCCCATCA 5842
 |||||
 RESULT 47
 AC108815 181547 bp DNA linear HTG 08-MAR-2003
 LOCUS
 DEFINITION Mus musculus clone RP23-239013, WORKING DRAFT SEQUENCE, 5 unordered
 pieces.
 ACCESSION AC108815
 VERSION AC108815.4 GI:28882391
 KEYWORDS HTG; HTGS, PHASE1; HTGS, DRAFT.
 SOURCE
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 181547)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-239013
 Unpublished
 TITLE
 REFERENCE 2 (bases 1 to 181547)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,

TITLE
JOURNALREFERENCE
AUTHORS

Direct Submission
 Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 181547)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
 Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (08-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 8, 2003 this sequence version replaced gi:28195423.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L20929
 Center clone name: 239_O_13
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 180713 bases at least Q40
 Consensus quality: 180940 bases at least Q30
 Consensus quality: 181059 bases at least Q20
 Insert size: 172000; agarose-fp
 Insert size: 181147; sum-of-contigs
 Quality coverage: 8.0 in Q20 bases; agarose-fp
 Quality coverage: 7.6 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces

TITLE
JOURNAL

COMMENT

4289 TTATTTCTGTAAGATGGTATTATACCTTCCTCCACCCCATCA 4332
 |||||
 5885 TTATTTCTGTAAGATGGTATTATACCTTCCTCCACCCCATCA 5842
 |||||
 RESULT 47
 AC108815 181547 bp DNA linear HTG 08-MAR-2003
 LOCUS
 DEFINITION Mus musculus clone RP23-239013, WORKING DRAFT SEQUENCE, 5 unordered
 pieces.
 ACCESSION AC108815
 VERSION AC108815.4 GI:28882391
 KEYWORDS HTG; HTGS, PHASE1; HTGS, DRAFT.
 SOURCE
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 181547)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-239013
 Unpublished
 TITLE
 REFERENCE 2 (bases 1 to 181547)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,

Tissue Procurement: Miklos Pavlovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Nakasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 76 Row: c Column: 5.

```

Series: max Place: 70 Max: 5 Column: 3
Location/Qualifiers
1. 505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44160 IMAGE:5287110"
/tissue_type="Brain, hypothalamus"

```

```

/notes=vector: bluescript"
23. .331
/codon_start=1
/product="unknown (protein for MGC:44160)"
/protein_id="AAH33953.1"
/db_xref="GI:21708023"
/translat_id="MCLF1FPFFFFFFFROSRLTSLPRLECSGTISAHCRRPGSCH
SPASASRVAGTGTGHHHTLWLCIFSRDGVSPCPGWSRSPDEVICLPQPKVLGIQA
"

```

```

      0.8%; Score 38; DB 9; Length 605;
h Similarity 100.0%; Pred.No. 6.4e-09;
38: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

8 TTTCTTTTTTCTTTTTTTTTTTTTTTCTTTTCCTTTTTTT 2405
9 TTTCTTTTTTCTTTTTTTTTTTTTTTCTTTTCCTTTTTTT 66

```

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143563)
Phillips.S
Direct Submission
Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

[illegible]

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., W.,
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, K., Hamilton, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Henderson, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kwio, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenschwartz, L., Loresaged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, B.,
Milosavljevic, A., Miner, G., Mirzaj, B., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Naif, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackaleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quirz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlczky, R., Woodden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 236171)
Worley, K.C.

Direct Submission
Submitted (10-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236171)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:22164899.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNWO
Center clone name: CH230-165K1
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 223310 bases at least Q40
Consensus quality: 224320 bases at least Q30
Consensus quality: 224905 bases at least Q20
Estimated insert size: 240553; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Ando, C., Dederich, D., Delgado, O., Denison, S., Dexamo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabis, A., Gante, R., Garcia, A., Garner, T., Garza, M., Geisregeorgis, E., Geir, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harunay, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowicz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Floresnshewa, L., Lozardo, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G.S., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., O'Nwalekeme, O., Okwuonu, G., Olarunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Platter, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Streimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 272182)
worcey, K.C.

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 272182)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23101339
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

PFMAL6P1_13
Continuation (6 of 14) of PFMAL6P1 from base 500001 -AL844505 Plasmodium falciparum 3D7
1300001 1378756

[illegible]


```

shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAOD
Center clone name: CH230-55G16
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 205676 bases at least Q40
Consensus quality: 209657 bases at least Q30
Consensus quality: 212140 bases at least Q20
Estimated insert size: 210246; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
** NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
** NOTE: This is a 'working draft' sequence. It currently
consists of 4 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
** This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
**
1 119045: contig of 119045 bp in length
**
1 119046 119145: gap of unknown length
**
1 119146 157294: contig of 38149 bp in length
**
1 157295 157394: gap of unknown length
**
1 157395 224562: contig of 67168 bp in length
**
224563 224662: gap of unknown length
**
224663 228253: contig of 1591 bp in length.
**
Location/Qualifiers
1 . 226253
/organism="Rattus norvegicus"
/mol type="genomic DNA"
/db xref="taxon:10116"
/clone="CH230-55G16"
**
1 . 2019
/notes="wgs_contig"
**
117096 . 119045
/notes="wgs contig"
**
188678 . 150943
/notes="wgs_contig"
**
h 0.7%; Score 35; DB 2; Length 226253;
Similarity 100.0%; Pred.No. 4.6e-07;
35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4 TTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTG 2408
|||||
4 TTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTG 7888
|||||
AC117916 227996 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-148E6, WORKING DRAFT SEQUENCE, 9
unordered pieces.
AC117916
AC117916.7 GI:25008861
HTG; HTGS PHASR1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 227996)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,

```

[illegible]

RESULT 64

AC094070

LOCUS

DEFINITION

AC094070

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 263236)

Muzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,B., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenschwa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangun,B., Mapua,P., Martin,K., Martin,R., Martine,E.,

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Miloavljjevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokelenah,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,

Puaio,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

Riley,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitrer,C.D., Smajd,D.,

Sneed,A., Sodergren,E., Song,X.-Z., Sorrellie,R., Sosa,J.,

Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,

Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,

Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von

Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 263236)

Worley,K.C.

Direct Submission

Submitted (14-SEP-2001)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 263236)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23321906.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GGUE

Center clone name: CH230-108N20

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 235442 bases at least Q40

Consensus quality: 239662 bases at least Q30

Consensus quality: 242686 bases at least Q20

Estimated insert size: 249310; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 261676: contig of 261676 bp in length

* 261677 261776: gap of unknown length

* 261777 263236: contig of 1460 bp in length.

FEATURES

source

1..263236

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-108N20"

1..1624

/note="wgs_end_extension"

clone_end:Sp6"

3854..4614

/note="clone boundary"

clone_end:Sp6

site:EcoRI

end sequence:BH278699"

259069..259952

/note="clone boundary"

clone_end:T7

site:EcoRI

end sequence:BH278696"

260192..261676

/note="wgs_end_extension"

clone_end:T7"

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 35; DB 2; Length 263236;

Db 111581 TTTTCCTTTTTTTTCTTTCTTTCTTTCTTTCTTTT 111615

RESULT 66
AX251462/c
LOCUS AX251462 14615 bp -DNA linear PAT 05-OCT-2001
DEFINITION Sequence 430 from Patent WO0168912.
ACCESSION AX251462
VERSION AX251462.1 GI:15984885

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.

TITLE Diagnosis of diseases associated with tumor suppressor genes and

oncogenes

JOURNAL Patent: WO 0168912-A, 430 20-SEP-2001;

FEATURES Epigenomics AG (DE)

source 1. 14615

Location/Qualifiers

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/notes="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 0.7%; Score 34; DB 6; Length 14615;

Best Local Similarity 100.0%; Pred. NO. 1.5e-06;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2374 TTTCTTTTTTTTTTTCTTTCTTTCTTTCTTTCTTTT 2407

Db 7601 TTTCTTTTTTTTTTTCTTTCTTTCTTTCTTTCTTTT 7568

Sequence split into 4 fragments LOCUS AC073763 Accession AC073763

Fragment Name Begin End

AC073763_0 1 110000

AC073763_1 100001 210000

AC073763_2 200001 310000

AC073763_3 300001 409770

LOCUS AC073763 409770 bp DNA linear HTG 29-JUN-2000

DEFINITION Mus musculus clone RP23-343P5, WORKING DRAFT SEQUENCE, 63 unordered

pieces.

ACCESSION AC073763

VERSION AC073763.1 GI:8810380

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 409770)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Mouse

Unpublished

REFERENCE 2 (bases 1 to 409770)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT -----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information

Center Project Name: 1868949

Center clone name: RPCI-23_343P5

Summary Statistics

Consensus quality: 35352 bases at least Q40
Consensus quality: 382147 bases at least Q30
Consensus quality: 388243 bases at least Q20
Estimated insert size: 217190; agarose-fp estimation
Estimated insert size: 403570; sum-of-contigs estimation
Quality coverage: 8.5 in Q20 bases; agarose-fp estimation
Quality coverage: 4.57 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1003: contig of 1003 bp in length
1004 1103: gap of unknown length
1104 2382: contig of 1279 bp in length
2383 2482: gap of unknown length
2483 3614: contig of 1132 bp in length
3615 3714: gap of unknown length
3715 4723: contig of 1009 bp in length
4724 4823: gap of unknown length
4823 6292: contig of 1469 bp in length
6293 7431: gap of unknown length
7432 7531: contig of 1039 bp in length
7532 8869: contig of 1338 bp in length
8870 9969: gap of unknown length
9970 10066: contig of 1097 bp in length
10067 10166: gap of unknown length
10167 11248: contig of 1082 bp in length
11249 11348: gap of unknown length
11349 13341: contig of 1993 bp in length
13342 13441: gap of unknown length
13442 14589: contig of 1148 bp in length
14590 14689: gap of unknown length
14690 15797: contig of 1108 bp in length
15798 15897: gap of unknown length
15898 16915: contig of 1018 bp in length
16916 17015: gap of unknown length
17016 18367: contig of 1352 bp in length
18368 18467: gap of unknown length
18468 19731: contig of 1264 bp in length
19732 19831: gap of unknown length
19832 20934: contig of 1103 bp in length
20935 21034: gap of unknown length
21035 22801: contig of 1767 bp in length
22802 22901: gap of unknown length
22902 24723: contig of 1822 bp in length
24724 24823: gap of unknown length
24824 26095: contig of 1272 bp in length
26096 26195: gap of unknown length
26196 27347: contig of 1152 bp in length
27348 27447: gap of unknown length
27448 28537: contig of 2090 bp in length
28538 29637: gap of unknown length
29638 31063: contig of 1426 bp in length
31064 31163: gap of unknown length
31164 33363: contig of 2200 bp in length
33364 33463: gap of unknown length
33464 36040: contig of 2577 bp in length
36041 36141: gap of unknown length
36141 38979: contig of 2839 bp in length
38980 39079: gap of unknown length
39080 42217: contig of 3138 bp in length
42218 42317: gap of unknown length
42318 44053: contig of 1736 bp in length
44054 44153: gap of unknown length
44154 45760: contig of 1607 bp in length
45761 45860: gap of unknown length
45861 48398: contig of 2538 bp in length

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.930319
Consensus quality: 128169 bases at least Q40
Consensus quality: 128674 bases at least Q30
Consensus quality: 128972 bases at least Q20
Insert size: 150000; agarose-1p
Quality coverage: 10.78x in Q20 bases; agarose-1p
Quality coverage: 12.52x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 10521: contig of 10521 bp in length
10522 10621: gap of unknown length
10622 13327: contig of 2706 bp in length
13328 13427: gap of unknown length
13428 28924: contig of 15497 bp in length
28925 29024: gap of unknown length
29025 39301: contig of 10177 bp in length
39302 41906: contig of 2605 bp in length
41907 42006: gap of unknown length
42007 70095: contig of 28089 bp in length
70096 70195: gap of unknown length
70196 77561: contig of 7366 bp in length
77562 94161: contig of 16500 bp in length
94162 97982: contig of 3721 bp in length
97983 98082: gap of unknown length
98083 105401: contig of 7318 bp in length
105401 105501: gap of unknown length
105501 127855: contig of 22355 bp in length
127856 127956: gap of unknown length
127956 130254: contig of 2299 bp in length.

FEATURES
source

Location/Qualifiers
1. .130254
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Brown Norway"
/db_xref="taxon:10116"
/clone_lib="RP31-258K6"
/clone_lib="RP31"

misc_feature

1. .10521
/notes="assembly_fragment"
clone_end:SP6
vector side:left

misc_feature

10622. .13327
/notes="assembly_fragment"

misc_feature

13428. 28924
/notes="assembly_fragment"

misc_feature

29025. .39201
/notes="assembly_fragment"

misc_feature

39302. .41906
/notes="assembly_fragment"

misc_feature

42007. .70095
/notes="assembly_fragment"

misc_feature

70196. .77561
/notes="assembly_fragment"

misc_feature

77662. .94161
/notes="assembly_fragment"

misc_feature

94262. .97982
/notes="assembly_fragment"

misc_feature

98083. .105400
/notes="assembly_fragment"

misc_feature

105501. .127855
/notes="assembly_fragment"

/note="assembly_fragment"

misc_feature

117804. .130254

/note="clone overlaps with GenBank Accession Number
AC134317 clone RP31-421B16 (center project name dce)"
127956. .130254

misc_feature

/note="assembly fragment
missing T7 clone end on 3' end of insert"

ORIGIN

Query Match 0.7%; Score 34; DB 2; Length 130254;

Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2374 TTTCCTTTTTTTTTTTCTTTTCTTTTTTTTTTTT 2407

DB 121459 TTTCCTTTTTTTTTTTCTTTTCTTTTTTTTTTTT 121492

RESULT 69

AC083783

LOCUS

AC083783 143516 bp DNA linear PRI 25-MAR-2003

Homo sapiens chromosome 17, clone RP11-462C21, complete sequence.

AC083783

DEFINITION

AC083783.8 GI:29171462

VERSION

HTG.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1. (bases 1 to 143516)

Birren, B., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 17, clone RP11-462C21

AUTHORS

Unpublished

JOURNAL

Unpublished

REFERENCE

2. (bases 1 to 143516)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,

Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Casale, P.,

Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,

Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,

FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,

Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,

Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K.,

Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G.,

Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,

McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,

Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,

O'Donnell, P., O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K.,

Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,

Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,

Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,

Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,

Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-SEP-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3. (bases 1 to 143516)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,

Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,

Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,

McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,

Mlenga, V., Morrow, J., Naylor, J., Norman, C., Nicol, R., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,

REFERENCE		FEATURES		Location/Qualifiers	
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/organism="Homo sapiens"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/db_type="genomic DNA"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/mol_xref="taxon:9606"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/chromosome="17"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/map="17"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/clone_lib="RPI1-462C21"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/clone_lib="RPI1-11 Human Male BAC"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluSg/x"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MIR"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluJo"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AT_rich"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluDb"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="FLAM_C"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MIR"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="(CA)n"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluSx"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="LTR33"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="FRAM"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluY"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="FLAM_C"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="LIMC4"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluDb"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="(TA)n"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="LIMB3A"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluDb"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MIR3"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="FAM"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="(CA)n"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MLT1K"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluY"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="LIMB3"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluSx"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AT_rich"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MIR"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="(10740..10867)"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="MER5B"
JOURNAL	AUTHORS	TITLE	COMMENT	repeated_region	/rpt_family="AluSc"

/note="clone overlaps with GenBank Accession Number

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	Direct Submission	Submitted (12-SEP-2002)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	Direct Submission	Submitted (12-SEP-2002)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
Barren, B., Nusbaum, C. and Lander, E.	Direct Submission	Submitted (12-SEP-2003)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Barren, B., Nusbaum, C. and Lander, E.	Direct Submission	Submitted (12-SEP-2003)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
Barren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menees, L., Minova, I., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, P., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	Direct Submission	Submitted (24-OCT-2003)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Barren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menees, L., Minova, I., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, P., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	Direct Submission	Submitted (24-OCT-2003)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:	Comment	On Oct 24, 2003	this sequence version replaced gi:34581718.	All repeats were identified using RepeatMasker:	Comment	On Oct 24, 2003	this sequence version replaced gi:34581718.
Smit, A. F. A. & Green, P. (1996-1997)	FTP	ftp://ftp.genome.washington.edu/RW/RepeatMasker.html		Smit, A. F. A. & Green, P. (1996-1997)	FTP	ftp://ftp.genome.washington.edu/RW/RepeatMasker.html	
Center: Whitehead Institute	Source	Center: Whitehead Institute	MIT Center for Genome Research	Center: Whitehead Institute	Source	Center: Whitehead Institute	MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu	Source	Web site: http://www-seq.wi.mit.edu		Web site: http://www-seq.wi.mit.edu	Source	Web site: http://www-seq.wi.mit.edu	
Contact: sequence_submissions@genome.wi.mit.edu	Source	Contact: sequence_submissions@genome.wi.mit.edu		Contact: sequence_submissions@genome.wi.mit.edu	Source	Contact: sequence_submissions@genome.wi.mit.edu	
Project Information	Source	Project Information		Project Information	Source	Project Information	
Center project name: L23683	Source	Center project name: L23683		Center project name: L23683	Source	Center project name: L23683	
Center clone name: 291_C4	Source	Center clone name: 291_C4		Center clone name: 291_C4	Source	Center clone name: 291_C4	
Location/Qualifiers	Source	Location/Qualifiers		Location/Qualifiers	Source	Location/Qualifiers	
1. 2918087	Source	1. 2918087		1. 2918087	Source	1. 2918087	
/organism="Mus musculus"	Source	/organism="Mus musculus"		/organism="Mus musculus"	Source	/organism="Mus musculus"	
/mol_type="genomic DNA"	Source	/mol_type="genomic DNA"		/mol_type="genomic DNA"	Source	/mol_type="genomic DNA"	
/db_xref="taxon:10090"	Source	/db_xref="taxon:10090"		/db_xref="taxon:10090"	Source	/db_xref="taxon:10090"	
/chromosome="9"	Source	/chromosome="9"		/chromosome="9"	Source	/chromosome="9"	
/map="9"	Source	/map="9"		/map="9"	Source	/map="9"	
/clone="RP23-291C4"	Source	/clone="RP23-291C4"		/clone="RP23-291C4"	Source	/clone="RP23-291C4"	
/clone.lib="RP23-291C4"	Source	/clone.lib="RP23-291C4"		/clone.lib="RP23-291C4"	Source	/clone.lib="RP23-291C4"	
/note="wgs end extension"	Source	/note="wgs end extension"		/note="wgs end extension"	Source	/note="wgs end extension"	
clone_end:SP6"	Source	clone_end:SP6"		clone_end:SP6"	Source	clone_end:SP6"	
147. 274	Source	147. 274		147. 274	Source	147. 274	
/rpt_family="(CA)n"	Source	/rpt_family="(CA)n"		/rpt_family="(CA)n"	Source	/rpt_family="(CA)n"	
281. 464	Source	281. 464		281. 464	Source	281. 464	
/rpt_family="WSTP"	Source	/rpt_family="WSTP"		/rpt_family="WSTP"	Source	/rpt_family="WSTP"	
1209. 1215	Source	1209. 1215		1209. 1215	Source	1209. 1215	
/note="<30 qual SNGL region"	Source	/note="<30 qual SNGL region"		/note="<30 qual SNGL region"	Source	/note="<30 qual SNGL region"	
1836. 1860	Source	1836. 1860		1836. 1860	Source	1836. 1860	
/rpt_family="(A)n"	Source	/rpt_family="(A)n"		/rpt_family="(A)n"	Source	/rpt_family="(A)n"	
complement(2878. 2972)	Source	complement(2878. 2972)		complement(2878. 2972)	Source	complement(2878. 2972)	
/rpt_family="PBID10"							

[illegible]

[illegible]

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 225253)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (04-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225253)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:25085370.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUPQ
Center clone name: CH230-195118
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 213689 bases at least Q40
Consensus quality: 215960 bases at least Q30
Consensus quality: 217430 bases at least Q20
Estimated insert size: 225330; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
* This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1 225253: contig of 225253 bp in length.
Location/Qualifiers
1 . 225253
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-195118"
1 . 1680
/note="wgs end_extension
clone end:17"
4312- 4881
/note="clone_boundary
clone end:17
site:EcoRI
end_sequence:BH344269"

AC102566
 AC102566.9 GI:39540561
 HTG.
 Mus musculus
 Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (bases 1 to 253343)
 Birren, B., Nusbaum, C. and Lander, E.

AUTHORS

TITLE
 Mus musculus chromosome 9, clone RP23-223D10

JOURNAL
 Unpublished

2 (bases 1 to 253343)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhalter, B.,
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, J., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
 3 (bases 1 to 253343)
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouckhalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cooke, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (09-NOV-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
 4 (bases 1 to 253343)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouckhalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,

Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (07-DEC-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
 On Dec 7, 2003 this sequence version replaced gi:38230045.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
 JOURNAL
 Direct Submission
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: W1BR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L19018
 Center clone name: 223_D_10
 ----- Location/Qualifiers
 1..253343
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /map="9"
 /clone="RP23-223D10"
 /clone_lib="RPCI-23 Female Mouse BAC"
 1..1232
 /note="wgs end extension
 clone end:SP6"
 repeat_region
 1..51
 /rpt_family="CT-rich"
 repeat_region
 129..340
 /rpt_family="MLT2F"
 repeat_region
 473..503
 /rpt_family="AT-rich"
 repeat_region
 527..829
 /rpt_family="MER2"
 repeat_region
 complement(1092..1230)
 /rpt_family="S1F"
 misc_feature
 1233..1238
 /note="clone boundary
 clone end:SP6
 site:EcoRI"
 repeat_region
 complement(1307..1838)
 /rpt_family="L1MDa"
 repeat_region
 2262..2297
 /rpt_family="RMR13B"
 repeat_region
 complement(3729..3914)
 /rpt_family="B2_Mm2"
 repeat_region
 complement(4397..4759)
 /rpt_family="L1MB1"
 repeat_region
 complement(4760..4917)
 /rpt_family="MTE"
 repeat_region
 4918..5326
 /rpt_family="MTD"
 repeat_region
 complement(5327..5526)
 /rpt_family="MTE"
 repeat_region
 complement(5528..5645)
 /rpt_family="L1MB1"
 repeat_region
 complement(5657..5801)
 /rpt_family="RMR1B"
 repeat_region
 complement(5802..6405)
 /rpt_family="Lx7"
 repeat_region
 complement(6411..6854)
 /rpt_family="BLTR10"

Accession	Contig	Length (bp)
4821	4820: contig of 4820 bp in length	4820
4921	4920: gap of unknown length	4920
4921	20160: contig of 15240 bp in length	15240
20161	20260: gap of unknown length	20260
20261	21819: contig of 3929 bp in length	3929
20261	24289: gap of unknown length	24289
24190	273663: contig of 248274 bp in length	248274
24190	273663: gap of unknown length	273663
272564	276688: contig of 4325 bp in length	4325
272664	277088: gap of unknown length	277088
276989	279386: contig of 1198 bp in length	1198
277089	279386: gap of unknown length	279386
278287	279310: contig of 1224 bp in length	1224
278387	279710: gap of unknown length	279710
279611	281063: contig of 1353 bp in length	1353
279711		

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwuo, S., Ogwu, M., Okwunonu, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojokan, J., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshkari, N., Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 384278)
Worley, K.C.

Direct Submission
Submitted (02-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 384278)
Worley, K.C.

Direct Submission
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 2, 2002 this sequence version replaced gi:22597262.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAKL
Center clone name: R11-911B23
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-Primer Bodypy: 5% of reads
Chemistry: Dye-terminator Big Dye: 95% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 339335 bases at least Q40
Consensus quality: 358089 bases at least Q30
Consensus quality: 371743 bases at least Q20
Estimated insert size: 354410; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
NOTE: This sequence may represent more than one 'clone'.
NOTE: This is a 'working draft' sequence. It currently consists of 42 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2053: contig of 2053 bp in length
2054 2153: gap of unknown length
2154 4452: contig of 2299 bp in length
4453 4552: gap of unknown length
4553 6757: contig of 2205 bp in length
6758 6857: gap of unknown length
6858 9164: contig of 2307 bp in length
9165 9264: gap of unknown length
9265 11882: contig of 2618 bp in length
11883 14493: gap of unknown length
14494 14544: gap of unknown length
14544 16762: contig of 2219 bp in length

16763 16862: gap of unknown length
16863 19467: contig of 2605 bp in length
19468 19567: gap of unknown length
19568 21657: contig of 2090 bp in length
21658 21757: gap of unknown length
21758 24842: contig of 3085 bp in length
24843 24942: gap of unknown length
24943 27517: contig of 2575 bp in length
27518 30503: contig of 2886 bp in length
30504 32613: gap of unknown length
32614 32714: gap of unknown length
32715 35628: contig of 2914 bp in length
35629 35728: gap of unknown length
35729 39450: contig of 3722 bp in length
39451 39550: gap of unknown length
39551 41621: contig of 2071 bp in length
41622 41721: gap of unknown length
41722 44132: contig of 2411 bp in length
44133 44232: gap of unknown length
44233 47404: contig of 3172 bp in length
47405 47504: gap of unknown length
47506 50765: contig of 3261 bp in length
50766 50865: gap of unknown length
50866 53511: contig of 2646 bp in length
53512 56573: contig of 2962 bp in length
56574 56782: contig of 2109 bp in length
56783 58882: gap of unknown length
58883 62123: contig of 3241 bp in length
62124 62223: gap of unknown length
62224 66245: contig of 4022 bp in length
66246 70164: contig of 3819 bp in length
70165 70264: gap of unknown length
70265 73265: contig of 3361 bp in length
73266 73725: gap of unknown length
73726 77079: contig of 3354 bp in length
77080 77179: gap of unknown length
77180 81163: contig of 3984 bp in length
81164 81263: gap of unknown length
81264 84305: contig of 3042 bp in length
84306 84405: gap of unknown length
84406 87928: contig of 3523 bp in length
87929 88028: gap of unknown length
88029 92621: contig of 4592 bp in length
92621 96922: gap of unknown length
96923 97022: contig of 4202 bp in length
97023 101484: contig of 4461 bp in length
101484 106242: contig of 4859 bp in length
106243 106342: gap of unknown length
106343 111035: contig of 4693 bp in length
111036 119325: contig of 8190 bp in length
119326 119425: gap of unknown length
119426 134873: contig of 15448 bp in length
134874 148356: contig of 13383 bp in length
148357 148456: gap of unknown length
148457 170779: contig of 22323 bp in length
170780 170879: gap of unknown length
170880 191306: contig of 20427 bp in length
191307 191406: gap of unknown length
191407 227798: contig of 36392 bp in length
227799 384278: gap of unknown length
384278 384278: contig of 156380 bp in length.

Location/Qualifiers
1. 384278
/organism="Homo sapiens"

FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-91E23"

ORIGIN
Query Match          0.7%; Score 33; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred.No. 6.2e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2369 TTTCTTTCCTTTTTTTTTTTCTTTTCTTTT 2401
Db       99019 TTTCTTTCCTTTTTTTTTTTCTTTTCTTTT 99051

RESULT 93
AL357518      115378 bp   DNA    linear   PRI 21-DEC-2000
LOCUS        Human DNA sequence from clone RP11-679B17 on chromosome 6, complete
DEFINITION   sequence.
ACCESSION    AL357518
VERSION      AL357518.15 GI:11967526
KEYWORDS     HTG.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 115378)
AUTHORS      Garner,P.
TITLE        Direct Submission
JOURNAL      Submitted (20-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
              requests: clonerquests@sanger.ac.uk
COMMENT      On Dec 22, 2000 this sequence version replaced gi:11863409.
              During sequence assembly data is compared from overlapping clones.
              Where differences are found these are annotated as variations
              together with a note of the overlapping clone name. Note that the
              variation annotation may not be found in the sequence submission
              corresponding to the overlapping clone, as we submit sequences with
              only a small overlap as described above.
              This sequence has been finished according to sequence map criteria
              as follows. An attempt is made to resolve all sequencing problems,
              such as compressions and repeats, but not necessarily within known
              annotated repeat sequence elements. Where the sequence is
              ambiguous there is an annotation using the 'unsure' feature key.
              The following abbreviations are used to associate primary accession
              numbers given in the feature table with their source databases:
              Em; ENBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPDB; Information
              on the WormPDB database can be found at
              http://www.sanger.ac.uk/projects/C_elegans/wormpdb This sequence
              was generated from part of bacterial clone contigs of human
              chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
              Group. Further information can be found at
              http://www.sanger.ac.uk/HGP/Chr6
              RP11-679B17 is from the library RPCI-11.3 constructed by the group
              of Pieter de Jong. For further details see
              http://www.chori.org/bacpac/home.htm
              VECTOR: pBACE3.6
              IMPORTANT: This sequence is not the entire insert of clone
              RP11-679B17 it may be shorter because we sequence overlapping
              sections only once, except for a 100 base overlap.
              The true left end of clone RP11-679B17 is at 1 in this sequence.
              The true left end of clone RP1-310P17 is at 115279 in this
              sequence. The true right end of clone RP11-716O23 is at 44457 in
              this sequence.

FEATURES             Location/Qualifiers
     1..115378
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"
         /chromosome="6"
         /clone="RP11-679B17"
         /clone_lib="RPCI-11"

```


O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ranasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (04-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 4, 2003 this sequence version replaced gi:11128427.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L107

Center clone name: 44_J_6

FEATURES

source

Location/Qualifiers

1..134457

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="CTB-44J6"

/clone_lib="CITB Human BAC"

1567..1672

/rpt_family="C-rich"

2044..2096

/rpt_family="CGGg/n"

2119..2154

/rpt_family="GC-rich"

2179..2242

/rpt_family="GC-rich"

2309..2332

/rpt_family="CA)n"

2661..2711

/rpt_family="CA)n"

3035..3057

/rpt_family="TC)n"

3057..3099

/rpt_family="CA)n"

complement(3835..4002)

/rpt_family="MER45A"

complement(5086..5237)

/rpt_family="MLTK"

complement(5236..5510)

/rpt_family="MLTK"

5600..5649

/rpt_family="TAGA)n"

complement(5691..5907)

/rpt_family="L2"

complement(5953..6124)

/rpt_family="L3"

6205..6371

/rpt_family="MIR"

complement(6367..6444)

/rpt_family="L3"

6470..6747

/rpt_family="L2"

7590..7738

/rpt_family="MIR3"

complement(8408..8530)

/rpt_family="L1MD3"

complement(8531..8819)

/rpt_family="AluSx"

complement(8820..9067)

repeat_region
complement(9943..10314)
/rpt_family="AluY"
repeat_region
complement(10627..10791)
/rpt_family="MIR"
11391..11657
/rpt_family="AluSg"
11658..11703
/rpt_family="CA)n"
11781..11837
/rpt_family="MIR"
11974..12101
/rpt_family="L2"
complement(13812..14103)
/rpt_family="AluSx"
14183..14284
/rpt_family="CATATA)n"
14198..14383
/rpt_family="CA)n"
15254..15420
/rpt_family="AluSg/x"
15768..15911
/rpt_family="MER45A"
16091..16300
/rpt_family="MIR"
16624..16684
/rpt_family="TC)n"
16736..17037
/rpt_family="AluSg"
complement(18237..18355)
/rpt_family="AluSg/x"
19358..19484
/rpt_family="L2"
complement(19485..19772)
/rpt_family="AluSg1"
19773..19993
/rpt_family="L2"
20176..20197
/rpt_family="AT-rich"
20287..20688
/rpt_family="L1M4"
20762..20894
/rpt_family="FLAM_C"
20944..20964
/rpt_family="AT-rich"
complement(21027..21410)
/rpt_family="MLT1B"
21475..21977
/rpt_family="L1MB4"
21980..22155
/rpt_family="TA)n"
22158..22482
/rpt_family="AluSx"
22803..22911
/rpt_family="L2"
23118..23241
/rpt_family="L3"
complement(23283..23588)
/rpt_family="AluY"
23755..23846
/rpt_family="L3"
24039..24233
/rpt_family="MER117"
24849..24861

Query Match 0.7% Score 33; DB:9; Length 134457;

Best Local Similarity 100.0%; Pred. No. 6.3e-06;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2372 TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2404

DB 28467 TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTT 28435

```

RESULT 95
AC130483/c
LOCUS      Mus musculus clone RP24-137D10, WORKING DRAFT SEQUENCE, 10
DEFINITION
  AC130483
  Mus musculus clone RP24-137D10, WORKING DRAFT SEQUENCE, 10
  unordered pieces.
ACCESSION
  AC130483
  GI:29294239
VERSION
  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  Birren,B., Nusbaum,C. and Lander,E.
  Mus musculus, clone RP24-137D10
  Unpublished
REFERENCE
  1 (bases 1 to 139684)
  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
  Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
  Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
  Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
  Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
  Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
  Karatas,A., Kells,C., Lander,E., Levine,R., Lindblad-Toh,K.,
  Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
  McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
  Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
  O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
  Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
  Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
  Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
  Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
  Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (10-AUG-2002) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  3 (bases 1 to 139684)
  Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
  Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
  Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
  Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
  Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
  Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
  Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
  Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
  Kamat,A., Karatas,A., Kells,C., Lander,E., Landers,T., Levine,R.,
  Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
  Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
  Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
  Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
  O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
  Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
  Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
  Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
  Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
  Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
  Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  On Mar 27, 2003 this sequence version replaced gi:28273447.
  All repeats were identified using RepeatMasker:
  Smit, A.F.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
  ----- Genome Center
  Center: Whitehead Institute/ MIT Center for Genome Research
  Center code: WIBR
  Web site: http://www-seq.wi.mit.edu
  Contact: sequence_submissions@genome.wi.mit.edu
  ----- Project Information

```

ORIGIN

```

Center project name: L27087
Center clone name: 137.D.10
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 137582 bases at least Q40
Consensus quality: 137980 bases at least Q30
Consensus quality: 138303 bases at least Q20
Insert size: 134000; agarose-fp
Insert size: 138784; sum-of-contigs
Quality coverage: 11.7 in Q20 bases; agarose-fp
Quality coverage: 11.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 8953: contig of 8953 bp in length
* 8854 8953: gap of 100 bp
* 8954 9621: contig of 668 bp in length
* 9622 9721: gap of 100 bp
* 9722 10403: contig of 682 bp in length
* 10404 10503: gap of 100 bp
* 10504 12335: contig of 2432 bp in length
* 12336 13035: gap of 100 bp
* 13036 16932: contig of 3797 bp in length
* 16933 35416: contig of 18484 bp in length
* 35417 35516: gap of 100 bp
* 35516 53015: contig of 17499 bp in length
* 53016 53115: gap of 100 bp
* 53116 76442: contig of 23327 bp in length
* 76443 76542: gap of 100 bp
* 76543 118440: contig of 41898 bp in length
* 118441 118540: gap of 100 bp
* 118541 139684: contig of 21144 bp in length.
*
* Location/Qualifiers
* 1..139684
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /clone="RP24-137D10"
  /clone_lib="RPCI-24 Male Mouse BAC"
*
* 1..8853
  /note="assembly_fragment"
  clone_end:8853
  vector_side:left
*
* 8954..9621
  /note="assembly_fragment"
  9722..10403
  /note="assembly_fragment"
  10504..12935
  /note="assembly_fragment"
  13036..16832
  /note="assembly_fragment"
  16933..35416
  /note="assembly_fragment"
  35517..53015
  /note="assembly_fragment"
  53116..76442
  /note="assembly_fragment"
  76543..118440
  /note="assembly_fragment"
  118541..139684
  /note="assembly_fragment"
  clone_end:17
  vector_side:right

```

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7/>, send email to: egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

This clone was derived from human PAC library RPc1-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com/>) or Research Genetics, Inc. (<http://www.resgen.com/>); or from Pieter de Jong; VECTOR: pCVPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is AC072049. Actual start of this clone is at base position 1 of RP5-1129J21 actual end is at base position 143773 of RP5-1129J21.

FEATURES

```

source
1. 143773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP5-1129J21"
/clone_lib="RPCI-5"
1. .72
/rpt_family="Alu"
repeat_region
111. .149
/rpt_family="Alu"
repeat_region
200. .498
/rpt_family="Alu"
repeat_region
499. .788
/rpt_family="Alu"
repeat_region
895. .1011
/rpt_family="MIR"
repeat_region
1315. .1613
/rpt_family="Alu"
repeat_region
2575. .2893
/rpt_family="Alu"
repeat_region
3194. .3502
/rpt_family="Alu"
repeat_region
3871. .3890
/rpt_family="(TTTAA)n"
repeat_region
3991. .4187
/rpt_family="Alu"
repeat_region
4201. .4363
/rpt_family="Alu"
repeat_region
4364. .4670
/rpt_family="Alu"
repeat_region
4671. .4807
/rpt_family="Alu"
repeat_region
4910. .4938
/rpt_family="Alu"
repeat_region
4991. .5296
/rpt_family="Alu"
repeat_region
6048. .6185
/rpt_family="Alu"

```

```

Query Match          0.78; Score 33; DB 2; Length 139684;
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2373  TTTTCTTTTTTTTTTTTTTTCTTTTCTTTTCTTTTCTTTT 2405
DDB      91735 TTTTCTTTTTTTTTTTTTTTCTTTTCTTTTCTTTTCTTTT 91703

RESULT 96
AC005631 LOCUS linear PRI 03-OCT-2003
DEFINITION Homo sapiens PAC clone RP5-112921 from 7, complete sequence.
AC005631 AC005631
VERSION 1
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143773)
Sulston, J.E. and Wilson, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 143773)
Bauer, C., Dante, M., Gregory, S., Hou, S. and Lamar, B.
The sequence of Homo sapiens PAC clone RP5-112921
Unpublished (2001)
JOURNAL
3 (bases 1 to 143773)
Waterston, R.H.
Direct Submission
Submitted (08-SEP-1998) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 143773)
Waterston, R.H.
Direct Submission
Submitted (03-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 143773)
Waterston, R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 143773)
Waterston, R.
Direct Submission
Submitted (26-APR-2003) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 143773)
Wilson, R.
Direct Submission
Submitted (03-OCT-2003) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 3, 2001 this sequence version replaced gi:13431072.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: saplens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H DJ1129J21
COMMENT

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

- * Runs of N , but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 12:45:34 ; Search time 315.852 Seconds
(without alignments)
8252.615 Million cell updates/sec

Title: US-10-023-523-17

Perfect score: 4697

Sequence: 1 caaaagcagccagcagca.....ttaaaaaaaaaaaaaaa 4697

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4697	100.0	4697	US-08-379-608A-17	Sequence 17, Appl
2	4697	100.0	4697	US-09-517-849-17	Sequence 17, Appl
3	4697	100.0	4697	US-09-616-289-17	Sequence 17, Appl
4	1439	30.6	1638	US-09-616-289-46	Sequence 46, Appl
5	844	18.0	22255	US-09-616-289-51	Sequence 51, Appl
6	774	16.5	1854	PCT-US94-01101-1	Sequence 1, Appl
7	218	4.6	492	US-09-833-381-1334	Sequence 1334, Ap
8	173	3.7	421	US-09-621-976-16512	Sequence 16512, A
9	59	1.3	4722	US-08-379-608A-14	Sequence 14, Appl
10	59	1.3	4722	US-09-517-849-14	Sequence 14, Appl
11	59	1.3	4722	US-09-616-289-14	Sequence 14, Appl
12	30	0.6	30	US-08-859-998-309	Sequence 309, App
13	30	0.6	30	US-08-859-998-310	Sequence 310, App
14	30	0.6	30	US-09-225-928-309	Sequence 309, App
15	30	0.6	30	US-09-225-928-310	Sequence 310, App
16	30	0.6	30	US-09-225-201B-309	Sequence 309, App
17	30	0.6	30	US-09-225-201B-310	Sequence 310, App
18	28	0.6	3001	US-09-539-333D-211	Sequence 211, Appl
19	27	0.6	953	US-09-508-824-17	Sequence 17, Appl
20	27	0.6	26664	US-09-564-805-28	Sequence 28, Appl
21	26	0.6	17000	US-09-548-797B-7	Sequence 7, Appl
22	26	0.6	78631	US-09-751-389-3	Sequence 3, Appl
23	25	0.5	224	US-08-731-272A-26	Sequence 26, Appl
24	25	0.5	4156	US-08-465-687A-1	Sequence 1, Appl
25	25	0.5	4156	US-09-030-970-1	Sequence 1, Appl
26	25	0.5	4156	US-09-520-210-1	Sequence 1, Appl
27	25	0.5	4156	PCT-US94-11843-1	Sequence 1, Appl

25	0.5	7970	1	US-08-135-511-31	Sequence 31, Appl
25	0.5	7997	1	US-08-187-453-31	Sequence 31, Appl
25	0.5	7997	1	US-08-562-985A-7	Sequence 7, Appl
25	0.5	14796	3	US-08-975-080-35	Sequence 35, Appl
25	0.5	14796	3	US-09-630-706-10	Sequence 10, Appl
25	0.5	14796	3	US-09-496-694B-3	Sequence 3, Appl
25	0.5	20303	1	US-08-370-975B-6	Sequence 6, Appl
25	0.5	26764	1	US-08-370-975B-1	Sequence 1, Appl
25	0.5	29604	3	US-08-781-891-207	Sequence 207, App
25	0.5	29604	4	US-09-618-166-207	Sequence 207, App
24	0.5	434	3	US-09-040-984-26	Sequence 26, Appl
24	0.5	434	4	US-09-123-912-26	Sequence 26, Appl
24	0.5	434	4	US-09-643-597-26	Sequence 26, Appl
24	0.5	434	4	US-09-480-884A-26	Sequence 26, Appl
24	0.5	434	4	US-09-542-615A-26	Sequence 26, Appl
24	0.5	434	4	US-09-606-421B-26	Sequence 26, Appl
24	0.5	434	4	US-09-221-107-26	Sequence 26, Appl
24	0.5	437	3	US-09-040-984-74	Sequence 74, Appl
24	0.5	437	4	US-03-123-912-74	Sequence 74, Appl
24	0.5	437	4	US-09-643-597-74	Sequence 74, Appl
24	0.5	437	4	US-09-480-884A-74	Sequence 74, Appl
24	0.5	437	4	US-09-542-615A-74	Sequence 74, Appl
24	0.5	437	4	US-09-606-421B-74	Sequence 74, Appl
24	0.5	437	4	US-09-221-107-74	Sequence 74, Appl
24	0.5	1080	1	US-07-885-970A-15	Sequence 15, Appl
24	0.5	1080	1	US-08-298-687A-15	Sequence 15, Appl
24	0.5	1080	1	US-08-530-797-16	Sequence 16, Appl
24	0.5	1080	1	US-08-298-829-15	Sequence 15, Appl
24	0.5	1080	2	US-08-787-335-16	Sequence 16, Appl
24	0.5	1272	4	US-09-673-395A-60	Sequence 60, Appl
24	0.5	2115	1	US-08-395-800A-7	Sequence 7, Appl
24	0.5	2540	4	US-08-684-708A-4	Sequence 4, Appl
24	0.5	3001	4	US-09-539-333D-195	Sequence 195, App
24	0.5	3517	3	US-08-420-306D-1	Sequence 1, Appl
24	0.5	4657	3	US-09-254-325-1	Sequence 1, Appl
24	0.5	12364	2	US-08-222-617A-1	Sequence 1, Appl
24	0.5	38653	4	US-09-932-445-1	Sequence 1, Appl
24	0.5	4453	4	US-09-146-053-5	Sequence 5, Appl
24	0.5	95500	4	US-09-798-096-10	Sequence 10, Appl
24	0.5	116592	4	US-08-818-512-3	Sequence 3, Appl
23	0.5	40	3	US-09-306-290-36	Sequence 36, Appl
23	0.5	225	4	US-09-702-705-612	Sequence 612, App
23	0.5	225	4	US-09-736-457-612	Sequence 612, App
23	0.5	225	4	US-09-614-124B-612	Sequence 612, App
23	0.5	225	4	US-09-671-325-612	Sequence 612, App
23	0.5	225	4	US-09-589-184-612	Sequence 612, App
23	0.5	260	2	US-08-520-678A-29	Sequence 29, Appl
23	0.5	260	3	US-08-897-126-29	Sequence 29, Appl
23	0.5	267	2	US-08-687-080-85	Sequence 85, Appl
23	0.5	519	4	US-09-621-976-1818	Sequence 1818, Ap
23	0.5	699	3	US-09-328-111-190	Sequence 190, App
23	0.5	925	4	US-03-257-179-17	Sequence 17, Appl
23	0.5	1288	1	US-08-047-041A-24	Sequence 24, Appl
23	0.5	1288	4	US-09-417-278A-1	Sequence 1, Appl
23	0.5	1316	1	US-08-047-041A-11	Sequence 11, Appl
23	0.5	1316	2	US-08-795-006A-31	Sequence 31, Appl
23	0.5	1316	3	US-09-184-073-31	Sequence 31, Appl
23	0.5	1974	1	US-08-413-135-3	Sequence 3, Appl
23	0.5	1974	4	US-08-971-395-3	Sequence 3, Appl
23	0.5	2065	4	US-03-277-196-3	Sequence 3, Appl
23	0.5	2066	4	US-09-023-655-1359	Sequence 1359, Ap
23	0.5	2646	1	US-08-539-304A-5	Sequence 5, Appl
23	0.5	2907	2	US-09-018-628-17	Sequence 17, Appl
23	0.5	2907	3	US-09-273-378-17	Sequence 17, Appl
23	0.5	2907	3	US-09-018-635-26	Sequence 26, Appl
23	0.5	2907	4	US-09-467-642-3	Sequence 3, Appl
23	0.5	2907	4	US-09-912-862-26	Sequence 26, Appl
23	0.5	3845	4	US-09-620-312D-554	Sequence 554, App
23	0.5	3845	4	US-09-620-312D-554	Sequence 554, App
23	0.5	4157	4	US-08-556-422A-1	Sequence 1, Appl
23	0.5	5228	4	US-10-164-595-23	Sequence 23, Appl
23	0.5	6124	4	US-08-213-419B-3	Sequence 3, Appl
23	0.5	8093	4	US-10-204-708-32	Sequence 32, Appl

Db 61 AGCGGCTCTGCTAGTAGAAGCAGAGGCTCCGGCAGCAGCCAGGCTCTCTCGAAGCCGG 120
Qy 121 AGGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGA 180
Db 121 AGGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGA 180
Qy 181 GCGCCAACTGGAAGACATACCTGAGCACAATCTGTGTGACAAATACACAGGGGGCCCGG 240
Db 181 GCGCCAACTGGAAGACATACCTGAGCACAATCTGTGTGACAAATACACAGGGGGCCCGG 240
Qy 241 GCGAGATGGGACAGAGGTGAGCGGCTGAACCCGGAAGATCAGAGAAGTCCCGGACCT 300
Db 241 GCGAGATGGGACAGAGGTGAGCGGCTGAACCCGGAAGATCAGAGAAGTCCCGGACCT 300
Qy 301 ATGTGCGAAGGATGGGGAGCTGAAACCACTCCAGTAGTCTATGAGAGAGGAGCCCT 360
Db 301 ATGTGCGAAGGATGGGGAGCTGAAACCACTCCAGTAGTCTATGAGAGAGGAGCCCT 360
Qy 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTCCAGAGTCCGAGAGCCGAGACC 420
Db 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTCCGAGAGTCCGAGAGCCGAGACC 420
Qy 421 ATCGAAGGCCACAG 480
Db 421 ATCGAAGGCCACAG 480
Qy 481 TGCAGACATTTGATCTCTGAGTACCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TGCAGACATTTGATCTCTGAGTACCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 ATGCTGAACCTGTGAG 600
Db 541 ATGCTGAACCTGTGAG 600
Qy 601 AGCAGAGCCAGCTGTGCAAG 660
Db 601 AGCAGAGCCAGCTGTGCAAG 660
Qy 661 TGSCCCCGCAGCAAGCTTGAG 720
Db 661 TGSCCCCGCAGCAAGCTTGAG 720
Qy 721 AGAAGAGGTGTGACAGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 AGAAGAGGTGTGACAGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 781 ACTTCAGAGTACACTGAATGATGATTCAGTGCAGATGAAACAGCAATGAGCGCAACT 840
Db 781 ACTTCAGAGTACACTGAATGATGATTCAGTGCAGATGAAACAGCAATGAGCGCAACT 840
Qy 841 CCAAGCTGGCCAG 900
Db 841 CCAAGCTGGCCAG 900
Qy 901 ATGAGCTGGCGAGGAG 960
Db 901 ATGAGCTGGCGAGGAG 960
Qy 961 TGTGTGATGCAAGAGCTCCAG 1020
Db 961 TGTGTGATGCAAGAGCTCCAG 1020
Qy 1021 AGCGGAG 1080
Db 1021 AGCGGAG 1080
Qy 1081 TGAAGCAG 1140
Db 1081 TGAAGCAG 1140
Qy 1141 AGTTCCAGAGACACACTTTCCAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 AGTTCCAGAGACACACTTTCCAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200

Qy 1201 AAAAGATGACTAAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 AAAAGATGACTAAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1261 GGGAGAGCAGCAACAAAG 1320
Db 1261 GGGAGAGCAGCAACAAAG 1320
Qy 1321 AACTGGAGGGGCTCAGAGTAAAAATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AACTGGAGGGGCTCAGAGTAAAAATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 CAGAGCCAGTACCTGAAACAG 1440
Db 1381 CAGAGCCAGTACCTGAAACAG 1440
Qy 1441 TCACTGACAGTGGGCTCAG 1500
Db 1441 TCACTGACAGTGGGCTCAG 1500
Qy 1501 CCAGGCTCAG 1560
Db 1501 CCAGGCTCAG 1560
Qy 1561 CTGGGCTCAG 1620
Db 1561 CTGGGCTCAG 1620
Qy 1621 GAAGGAGCGCAG 1680
Db 1621 GAAGGAGCGCAG 1680
Qy 1681 TTGCTGAAGCCAG 1740
Db 1681 TTGCTGAAGCCAG 1740
Qy 1741 GTGGTCAAGTGTTCAGTACATAG 1800
Db 1741 GTGGTCAAGTGTTCAGTACATAG 1800
Qy 1801 TAAGTGTACAGTGGGCTGAG 1860
Db 1801 TAAGTGTACAGTGGGCTGAG 1860
Qy 1861 TCTGTAGCTAAGAGAGTCTTGAAG 1920
Db 1861 TCTGTAGCTAAGAGAGTCTTGAAG 1920
Qy 1921 GCAGAGTACCTGAG 1980
Db 1921 GCAGAGTACCTGAG 1980
Qy 1981 CAGAGCTCAAAACAG 2040
Db 1981 CAGAGCTCAAAACAG 2040
Qy 2041 TGTGTGTGAGTCAAG 2100
Db 2041 TGTGTGTGAGTCAAG 2100
Qy 2101 AAAAGTGTCTTAGGAG 2160
Db 2101 AAAAGTGTCTTAGGAG 2160
Qy 2161 ACAAGGAG 2220
Db 2161 ACAAGGAG 2220
Qy 2221 ATCGCTTCCCTGAG 2280
Db 2221 ATCGCTTCCCTGAG 2280

QY		2281	TATCAITTCACAGGTGCCTCTCTACGCTGCGCAAAATGATTGACAAGAGATCACCCAAAGGA	2340
Db		2281	TATCAITTCACAGGTGCCTCTCTACGCTGCGCAAAATGATTGACAAGAGATCACCCAAAGGA	2340
QY		2341	TTATTTCTGGAAGGTGTTTTTCTTTTATTAATCTTTTCTTTTTTTTTTTTTTTTTTTTTCTTTT	2400
Db		2341	TTATTTCTGGAAGGTGTTTTTCTTTTATTAATCTTTTCTTTTTTTTTTTTTTTTTTTTTCTTTT	2400
QY		2401	TTTTTTTTGCACATGCACAGTGTTTTGTATTGAGGACCTTCCAAGGAAAAGGATGCTGTACC	2460
Db		2401	TTTTTTTTGCACATGCACAGTGTTTTGTATTGAGGACCTTCCAAGGAAAAGGATGCTGTACC	2460
QY		2461	AGTGGTGCTGGGTGCCTGCGCTCCAGTGTCGCCACTCCTTCCACCAACCCCATTTGGCTCC	2520
Db		2461	AGTGGTGCTGGGTGCCTGCGCTCCAGTGTCGCCACTCCTTCCACCAACCCCATTTGGCTCC	2520
QY		2521	TTTGCCATCTTGATGCTGAGGTTTTCTCTGTTGTGTGAGATCAGGTGTGTTGTGTAATAAGA	2580
Db		2521	TTTGCCATCTTGATGCTGAGGTTTTCTCTGTTGTGTGAGATCAGGTGTGTTGTGTAATAAGA	2580
QY		2581	AAGAAAGGGCTTCTGATGCTTTGGCAAAGCTTACCTGTGGGTTTCAGTCTGAGAGG	2640
Db		2581	AAGAAAGGGCTTCTGATGCTTTGGCAAAGCTTACCTGTGGGTTTCAGTCTGAGAGG	2640
QY		2641	CCACCACAGTCCCATCAGCACTGTCTCCATGCAGCAGTGTCTGGGTCCCATGTCCAGC	2700
Db		2641	CCACCACAGTCCCATCAGCACTGTCTCCATGCAGCAGTGTCTGGGTCCCATGTCCAGC	2700
QY		2701	TGCTCTTTGGCTTCATGGGTTTTTCTGCTTCCTGCCCCCACCCACATGTGCAATCCT	2760
Db		2701	TGCTCTTTGGCTTCATGGGTTTTTCTGCTTCCTGCCCCCACCCACATGTGCAATCCT	2760
QY		2761	CAGATTTGCTCGATTCTAATTTCTGGCACCTCCCTGCTGTCTGGGGATCTCACTT	2820
Db		2761	CAGATTTGCTCGATTCTAATTTCTGGCACCTCCCTGCTGTCTGGGGATCTCACTT	2820
QY		2821	CTTCTGTGTGGGGCCCATAGCTGTTGTCTAACAGGTAAAGAAATGAAATTGAACTATTGA	2880
Db		2821	CTTCTGTGTGGGGCCCATAGCTGTTGTCTAACAGGTAAAGAAATGAAATTGAACTATTGA	2880
QY		2881	CTGGGCCCCAGAAATCCATAAAATGGCTGCAGACAGTGTGTTCTGTCTGTCTTACCC	2940
Db		2881	CTGGGCCCCAGAAATCCATAAAATGGCTGCAGACAGTGTGTTCTGTCTGTCTTACCC	2940
QY		2941	CCACTCCAGTACATAACTACTATGTAAGTAGAGCCATTCTATATGCTGAATGTTCTG	3000
Db		2941	CCACTCCAGTACATAACTACTATGTAAGTAGAGCCATTCTATATGCTGAATGTTCTG	3000
QY		3001	CTGTGCAAACTTGCCAGGGTATTAGCCAGTGTGTCGCAACAGGTTTTTCGGGGACAAC	3060
Db		3001	CTGTGCAAACTTGCCAGGGTATTAGCCAGTGTGTCGCAACAGGTTTTTCGGGGACAAC	3060
QY		3061	AGAAATGACTCAGACCAAGATGGATAGGTTAGGCTTTGCTTCTGTGTTTCTTCTT	3120
Db		3061	AGAAATGACTCAGACCAAGATGGATAGGTTAGGCTTTGCTTCTGTGTTTCTTCTT	3120
QY		3121	TGAACCTAGTCAATTTGCTGCGAGTCCCTTCANCTTCCATACCTAGCCCACTTTTTAGCC	3180
Db		3121	TGAACCTAGTCAATTTGCTGCGAGTCCCTTCANCTTCCATACCTAGCCCACTTTTTAGCC	3180
QY		3181	CTTACCTTTAAATCTCTCAGATAAGTTGGTTTCAAAAAGATGTTAACTAGTCAATCATGTG	3240
Db		3181	CTTACCTTTAAATCTCTCAGATAAGTTGGTTTCAAAAAGATGTTAACTAGTCAATCATGTG	3240
QY		3241	TGACTGAGACGAGATGGCAAAATGAATGSCACACCAATTTCTCCTTCTCTGCCCCAGGG	3300
Db		3241	TGACTGAGACGAGATGGCAAAATGAATGSCACACCAATTTCTCCTTCTCTGCCCCAGGG	3300
QY		3301	CAGGTACCACTGATCTGCATCAGAGTTGCTGCTATTCTCTGCTGTATCTCTTCACTA	3360
Db		3301	CAGGTACCACTGATCTGCATCAGAGTTGCTGCTATTCTCTGCTGTATCTCTTCACTA	3360
QY		3361	GGTGCCCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATACCTC	3420

3361	GGTGCCTTCAAGCAGCTGTGTGAGTGTTCAGATCTCTGCCATCTCTGGCTCAGATATGCG	3420
3421	TGCTCTGTGAAGTGTTCCTCATCAGACCTTTTCTTCCCTTTGAAATCCCTCTTGTCTGGAG	3480
3421	TGTCTCTGTGAAGTGTTCCTCATCAGACCTTTTCTTCCCTTTGAAATCCCTCTTGTCTGGAG	3480
3481	TAGTCCCTTGCCTTCTTCTTGCTCCAGTAGGCCCTTTTCTTACCCAGGCCCTTGTGCCAGG	3540
3481	TAGTCCCTTGCCTTCTTCTTGCTCCAGTAGGCCCTTTTCTTACCCAGGCCCTTGTGCCAGG	3540
3541	CTAAGCTGTACAAGAGCTGCCAACTCACAGAGTTTGTCTAGGCGAGAGGTCAGAGGA	3600
3541	CTAAGCTGTACAAGAGCTGCCAACTCACAGAGTTTGTCTAGGCGAGAGGTCAGAGGA	3600
3601	AGAGCAGAGGTATGCACCTTTCCTCCCTTGAAGAGAGGGGAAAGGCTTACAGTGGCCCCACA	3660
3601	AGAGCAGAGGTATGCACCTTTCCTCCCTTGAAGAGAGGGGAAAGGCTTACAGTGGCCCCACA	3660
3661	TAAATTGCTGACTCACACTTTCAGCTACCTCTTAATGCTGCTGTGAGGAGACTGGAGCTGCTG	3720
3661	TAAATTGCTGACTCACACTTTCAGCTACCTCTTAATGCTGCTGTGAGGAGACTGGAGCTGCTG	3720
3721	GATCCAGTGTGTGTGTAGGAGGCCACAGTGAGCAGGTGGCCCCAGCTGGTGTTCCTCA	3780
3721	GATCCAGTGTGTGTGTAGGAGGCCACAGTGAGCAGGTGGCCCCAGCTGGTGTTCCTCA	3780
3781	GGTCAAGAAATGTGGCCCCCAGGCAAGGTGAGGCTTTGCTCACAGCTCCATCAATGTCCTA	3840
3781	GGTCAAGAAATGTGGCCCCCAGGCAAGGTGAGGCTTTGCTCACAGCTCCATCAATGTCCTA	3840
3841	GACCTTCAGGCCAGCTGTGCAGATGAGGTTCCCTACCTTTTCTTCTCATTCAGCACCAAA	3900
3841	GACCTTCAGGCCAGCTGTGCAGATGAGGTTCCCTACCTTTTCTTCTCATTCAGCACCAAA	3900
3901	TCACCAATACCTACAGCTGCTCTGCTTCTGCTTCCAAAGTACCCAGGTCCTGGGCCA	3960
3901	TCACCAATACCTACAGCTGCTCTGCTTCTGCTTCCAAAGTACCCAGGTCCTGGGCCA	3960
3961	GATGCAGGGAGGTGCTCTATCATGAGTCAAGGCCAGTGTCTTCTCATCCTGGTGGTCTCC	4020
3961	GATGCAGGGAGGTGCTCTATCATGAGTCAAGGCCAGTGTCTTCTCATCCTGGTGGTCTCC	4020
4021	CACACTTGTGACCTCAGTTTATAGGACCCAAAGATCTGTGTGGTTCCTTAGATGCTTAGC	4080
4021	CACACTTGTGACCTCAGTTTATAGGACCCAAAGATCTGTGTGGTTCCTTAGATGCTTAGC	4080
4081	TTTTTCTCCAGGGACACACAGCAGGTGAGCTCAGAGGCCATGGCTCTGCTAATAGTAA	4140
4081	TTTTTCTCCAGGGACACACAGCAGGTGAGCTCAGAGGCCATGGCTCTGCTAATAGTAA	4140
4141	ATTGTGTTTCAGGGCCTTGTCCAGCTCAGAGCTTCATGTCCACCAAGTCTGAGAGGTGTC	4200
4141	ATTGTGTTTCAGGGCCTTGTCCAGCTCAGAGCTTCATGTCCACCAAGTCTGAGAGGTGTC	4200
4201	AGCAGCACTTTTATTTTATTTGTTGTTTTCATGAGGTATCGGACATGGGCT	4260
4201	AGCAGCACTTTTATTTTATTTGTTGTTTTCATGAGGTATCGGACATGGGCT	4260
4261	GAGCTCAGGCACCTTCTGTAGGAGACTGTTATTTCTGTAAAGATGGTTATTTAACCCCTCC	4320
4261	GAGCTCAGGCACCTTCTGTAGGAGACTGTTATTTCTGTAAAGATGGTTATTTAACCCCTCC	4320
4321	TCACCCCATCA CGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCCCTGGTGTC	4380
4321	TCACCCCATCA CGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCCCTGGTGTC	4380
4381	CACGGGGAGGGCCAGAGGCTGCTGAGCTGATTCTTCAGCTGCTGCCCGCAGCCTTCCCG	4440
4381	CACGGGGAGGGCCAGAGGCTGCTGAGCTGATTCTTCAGCTGCTGCCCGCAGCCTTCCCG	4440
4441	CTTGCA GAGCAGAGGTTGTACCCCGAGGACAGCCAGGCACTGTCTCTTGGCCCTT	4500

Db 4441 CTTGCACAGCAGAGGTGTCACCCAGGACAGCCAGGACCTGCTCTCTGCCCC 4500
QY 4501 CTTGGGGGAAAGAGCTGCTCTGTCCTGTAACTGCTTCTTATGCCCCAAACCCGCGC 4560
Db 4501 CTTGGGGGAAAGAGCTGCTCTGTCCTGTAACTGCTTCTTATGCCCCAAACCCGCGC 4560
QY 4561 CACTCAGACTTGTTCGAAGCTGCACATGGCAGCTTTTGTCTCTCTTGGGTATTCACAAAC 4620
Db 4561 CACTCAGACTTGTTCGAAGCTGCACATGGCAGCTTTTGTCTCTCTTGGGTATTCACAAAC 4620
QY 4621 AGCCAGGACTTGATTTTGATGATTTTAAACACACATTAATAAAGAGTCTGTTGCCCTTA 4680
Db 4621 AGCCAGGACTTGATTTTGATGATTTTAAACACACATTAATAAAGAGTCTGTTGCCCTTA 4680
QY 4681 AAAAAAAAAAAAAA 4697
Db 4681 AAAAAAAAAAAAAA 4697

RESULT 3
US-09-616-289-17
; Sequence 17, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17

Query Match 100.0%; Score 4697; DB 4; Length 4697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAAAGCAGCCAGGACAAACCGGAGCAGGACCCGAGGAGCCCAAGGAGCCGCGCCAGCC 60
Db 1 CAAAAAGCAGCCAGGACAAACCGGAGCAGGACCCGAGGAGCCCAAGGAGCCGCGCCAGCC 60
QY 61 AGCGGCTCTCTGAGTAGAGCAGAGAGTCCCGGAGCAGGAGCTCTCTCGAGAGCCG 120
Db 61 AGCGGCTCTCTGAGTAGAGCAGAGAGTCCCGGAGCAGGAGCTCTCTCGAGAGCCG 120
QY 121 AGGGGGCTCAAGCAGAACCGGCTCAGTCTGGGCCCCCTTGATGATCTCTGAGGAGCTGA 180
Db 121 AGGGGGCTCAAGCAGAACCGGCTCAGTCTGGGCCCCCTTGATGATCTCTGAGGAGCTGA 180
QY 181 GCGCCCACTGAGACATCTGAGCACAATCTGTGGACATATACCAACAGGGGGGCCCCG 240
Db 181 GCGCCCACTGAGACATCTGAGCACAATCTGTGGACATATACCAACAGGGGGGCCCCG 240

QY 241 GCGAGGATGGGCGCAGAGGTGAGCCGCTGAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 300
Db 241 GCGAGGATGGGCGCAGAGGTGAGCCGCTGAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 300
QY 301 ATGTGGCAAGGAATGGGAGGAGCTGAACCAACTCCAGTAGTCTATGGAGAGAGAAACCT 360
Db 301 ATGTGGCAAGGAATGGGAGGAGCTGAACCAACTCCAGTAGTCTATGGAGAGAGAAACCT 360
QY 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGAGCCGAGACC 420
Db 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGAGCCGAGACC 420
QY 421 ATCAAGGCCACAGAGAGAGAAACCAAGGGTTTGGGAGAGAGATCACGTTGTCTGA 480
Db 421 ATCAAGGCCACAGAGAGAGAAACCAAGGGTTTGGGAGAGAGATCACGTTGTCTGA 480
QY 481 TGCAGACATTGAATTAATCTGAGTACCCAGAGGAGAGCTGGCTGCTCTGTGCAAGAGT 540
Db 481 TGCAGACATTGAATTAATCTGAGTACCCAGAGGAGAGCTGGCTGCTCTGTGCAAGAGT 540
QY 541 ATGCTGAATCTCTGAGGAGCACCAGGAATTCACAGAGCAGATGAAGCTCCTACAGAAA 600
Db 541 ATGCTGAATCTCTGAGGAGCACCAGGAATTCACAGAGCAGATGAAGCTCCTACAGAAA 600
QY 601 AGCAGAGCCAGCTGGTCAAGAGAGAGACCACTTGCAGGAGTGCAGCAGCAGCAAGCCGCTCC 660
Db 601 AGCAGAGCCAGCTGGTCAAGAGAGAGACCACTTGCAGGAGTGCAGCAGCAGCAAGCCGCTCC 660
QY 661 TGGCCCGCAGCAAGCTTGAGAGCCTATGCCCTGAGCTGCAGCGGCAACAACCGCTCCCTCA 720
Db 661 TGGCCCGCAGCAAGCTTGAGAGCCTATGCCCTGAGCTGCAGCGGCAACAACCGCTCCCTCA 720
QY 721 AGGAAGAGTGTGTCAGCGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 AGGAAGAGTGTGTCAGCGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 ACTTCCAGTGCACACTGATTCAGCTGAGATGGAACAGCAGCAATCAGCGCACT 840
Db 781 ACTTCCAGTGCACACTGATTCAGCTGAGATGGAACAGCAGCAATCAGCGCACT 840
QY 841 CCAAGCTCGCCCAAGAGAACATGAGGCTGGTGCAGAGGCTCAAGAAGCTGATTGAGCAGT 900
Db 841 CCAAGCTCGCCCAAGAGAACATGAGGCTGGTGCAGAGGCTCAAGAAGCTGATTGAGCAGT 900
QY 901 ATGAGCTCGGAGAGAGCATATCGCAAAAGTCTTCAACACAGAGGAGCTTACAGCAGC 960
Db 901 ATGAGCTCGGAGAGAGCATATCGCAAAAGTCTTCAACACAGAGGAGCTTACAGCAGC 960
QY 961 TGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAGC 1020
Db 961 TGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAGC 1020
QY 1021 AGCGGAGAGAGATTTCTCTGAAAGAGGAGTGCAGTCCAGAGGATGTGTGAGCTGA 1080
Db 1021 AGCGGAGAGAGATTTCTCTGAAAGAGGAGTGCAGTCCAGAGGATGTGTGAGCTGA 1080
QY 1081 TGAACGACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1081 TGAACGACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1141 AGTTCCAGAACACATTTTCCAAAAGCAGGAGTATTCACCAATTCAGCAGGAGATGG 1200
Db 1141 AGTTCCAGAACACATTTTCCAAAAGCAGGAGTATTCACCAATTCAGCAGGAGATGG 1200
QY 1201 AAAAGATCACTAAGAGATCAAGAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db 1201 AAAAGATCACTAAGAGATCAAGAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 GGGAGAGCAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
Db 1261 GGGAGAGCAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
QY 1321 AACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380

1321 AACCTGGAGGCGCTGCAGGTAAATCAACGGCTGGAGAGCTGTCCGGGCACTGCAGA 1380
1381 CAGAGCGCAATGACCTGAACAGAGGGTACAGAGCTGAGTGTGTGGTGGCCAGGGCTCCC 1440
1381 CAGAGCGCAATGACCTGAACAGAGGGTACAGAGCTGAGTGTGTGGTGGCCAGGGCTCCC 1440
1441 TCACCTGACAGTGGCCCTGAGAGGAGGCCAGAGGGGCTGGGGCTCAAGACCCAGCTCCC 1500
1441 TCACCTGACAGTGGCCCTGAGAGGAGGCCAGAGGGGCTGGGGCTCAAGACCCAGCTCCC 1500
1501 CCAGGGTCA CAGAGCGCCCTTGCTACCCAGGAGCA CCGAGCA CAGAGCAATCAGGCCAGA 1560
1501 CCAGGGTCA CAGAGCGCCCTTGCTACCCAGGAGCA CCGAGCA CAGAGCAATCAGGCCAGA 1560
1561 CTGGGCTCAAGAGCCACCTCCGCCAGAGGCTTAGAGAGCTGGTGTGGTCACTCTGG 1620
1561 CTGGGCTCAAGAGCCACCTCCGCCAGGCTTAGAGAGCTGGTGTGGTCACTCTGG 1620
1621 GAAGGAGCGGCA GCGCCAGCGAGGCTGSCCCATAAAGGCTCCCATGTGAGCAGCCCA 1680
1621 GAAGGAGCGGCA GCGCCAGCGAGGCTGSCCCATAAAGGCTCCCATGTGAGCAGCCCA 1680
1681 TTGCTGAAGCCAGGATGTTCTTGAACCTGGCTGGCATCTGGCACTTGCATTTGGATTTT 1740
1681 TTGCTGAAGCCAGGATGTTCTTGAACCTGGCTGGCATCTGGCACTTGCATTTGGATTTT 1740
1741 GTGGCTCAGTATTACGTACATAGGCAATTTTGAAGGCTTGCATTTGATTTATACCTG 1800
1741 GTGGCTCAGTATTACGTACATAGGCAATTTTGAAGGCTTGCATTTGATTTATACCTG 1800
1801 TAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
1801 TAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGTGTGTACAGATGAAGTCAGTGGCTTG 1860
1861 TCTGTGAGCTGAAGAGTCTTGAAGGGGCTGTCTATCTGTAGCTGCCATCACAGTGAAGTTG 1920
1861 TCTGTGAGCTGAAGAGTCTTGAAGGGGCTGTCTATCTGTAGCTGCCATCACAGTGAAGTTG 1920
1921 GCAGAGTGAATTGAGCAATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980
1921 GCAGAGTGAATTGAGCAATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT 1980
1981 CAGAGCTCAAAACAGTAATACACAGGCTTGTGATTTGCTGTGAGCAGGGCT 2040
1981 CAGAGCTCAAAACAGTAATACACAGGCTTGTGATTTGCTGTGAGCAGGGCT 2040
2041 TGCTTGTGAGCTCAGGCGCTCCTAGCTGTGTGGAGGCTCCTTTGATTTCTTAGACCTGG 2100
2041 TGCTTGTGAGCTCAGGCGCTCCTAGCTGTGTGGAGGCTCCTTTGATTTCTTAGACCTGG 2100
2101 AAAAGGTGCTCCCTAGGAGAGCCCTGGCAGGCGCTCAGAGCTGGATTTCTGCTGGA 2160
2101 AAAAGGTGCTCCCTAGGAGAGCCCTGGCAGGCGCTCAGAGCTGGATTTCTGCTGGA 2160
2161 ACAAGGACCTGAGAGATGTTTTTGGTGGATGATGTGTGTGAGGAGCCCTTGGGC 2220
2161 ACAAGGACCTGAGAGATGTTTTTGGTGGATGATGTGTGTGAGGAGCCCTTGGGC 2220
2221 ATGCTTCCCTCCCTTTGGTGTGAGGAGCAGGAGCCAGGCAATGATGCTTCTAGTAGCTT 2280
2221 ATGCTTCCCTCCCTTTGGTGTGAGGAGCAGGAGCCAGGCAATGATGCTTCTAGTAGCTT 2280
2281 TATCATTTACAGGTGCTCTAGCTGTGCAAAATGATTTGACAGAGATCACCCAAAGGA 2340
2281 TATCATTTACAGGTGCTCTAGCTGTGCAAAATGATTTGACAGAGATCACCCAAAGGA 2340
2341 TTAATTTCTGAAGGTGTTTTTTCTTAATTTCTTTTTTTCTTTTTTTCTTTTTTTCTTTT 2400
2341 TTAATTTCTGAAGGTGTTTTTTCTTAATTTCTTTTTTTCTTTTTTTCTTTTTTTCTTTT 2400
2401 TTTTTTTGACATGACAGTGTGTGATTTGAGGACCTTCAAGGAAAGGATGCTGTACC 2460

2401 TTTTTTTGACATGACAGTGTGTGTGATTTGAGGACCTTCAAGGAAAGGGATGCTGTACC 2460
2461 AGTGTGCTCGGGTGGCTTGGCTCCAGTGTCCACACTCCTTCCACACCCCACTTGGGTCC 2520
2461 AGTGTGCTCGGGTGGCTTGGCTCCAGTGTCCACACTCCTTCCACACCCCACTTGGGTCC 2520
2521 TTTGCCATTTGATGCTGAGGCTTCTGTGTTGGTGTGAGATCAGGTTGTTGGTGTAAAGA 2580
2521 TTTGCCATTTGATGCTGAGGCTTCTGTGTTGGTGTGAGATCAGGTTGTTGGTGTAAAGA 2580
2581 AAGGAAAGGGCTTCTGATGGCTTGGCACAAGCTTACTGTGGGTTTCAGTCTCCAGG 2640
2581 AAGGAAAGGGCTTCTGATGGCTTGGCACAAGCTTACTGTGGGTTTCAGTCTCCAGG 2640
2641 CCACACAGTTCCTCCATCAGCACTGTCTCCATGAGCAGTGTGTGGTCCCATGTCCAGC 2700
2641 CCACACAGTTCCTCCATCAGCACTGTCTCCATGAGCAGTGTGTGGTCCCATGTCCAGC 2700
2701 TGCTCTTTGGCTTCATGGGTTTTCCTGCTTCCCTGCCCCACCCCCACATGTGCAATCCT 2760
2701 TGCTCTTTGGCTTCATGGGTTTTCCTGCTTCCCTGCCCCACCCCCACATGTGCAATCCT 2760
2761 CAAGATTTGCTCCTGATTTCTATTTCTGCGCACCTCCCTGCGCTGCTTGGGATTTACTT 2820
2761 CAAGATTTGCTCCTGATTTCTATTTCTGCGCACCTCCCTGCGCTGCTTGGGATTTACTT 2820
2821 CTTCCTGTGTGGGCCCCATAGCTGTTGTCTAAAGGTAAAGAAATGAAATTAATATGA 2880
2821 CTTCCTGTGTGGGCCCCATAGCTGTTGTCTAAAGGTAAAGAAATGAAATTAATATGA 2880
2881 CTGGGCCCCCAAAATCCATAAAATGGCTGAGCAGTGTGTTCTGTCTCTTCTTACCC 2940
2881 CTGGGCCCCCAAAATCCATAAAATGGCTGAGCAGTGTGTTCTGTCTCTTCTTACCC 2940
2941 CCACCTCAGTACATAAATCTACTATCTGTGTAGAGCCATTTCTATATGCTGAATGTTCTG 3000
2941 CCACCTCAGTACATAAATCTACTATCTGTGTAGAGCCATTTCTATATGCTGAATGTTCTG 3000
3001 CTGTGTGAACTTCCAGGCTTATAGCAGTGTGTTGGCCAGGATTTTGGGAGCAAC 3060
3001 CTGTGTGAACTTCCAGGCTTATAGCAGTGTGTTGGCCAGGATTTTGGGAGCAAC 3060
3061 AGAATGACTCAGACCAAGATGATAGGATGTTAGGCTTGTCTTCTTGTGTTTTCTT 3120
3061 AGAATGACTCAGACCAAGATGATAGGATGTTAGGCTTGTCTTCTTGTGTTTTCTT 3120
3121 TGAATGACTGATTTCTGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3180
3121 TGAATGACTGATTTCTGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3180
3181 CTATCTTAAATCTCTCAGATAAGTGTGTTTCAAAAGAAATGTTTAAAGTACTGAATCATGTG 3240
3181 CTATCTTAAATCTCTCAGATAAGTGTGTTTCAAAAGAAATGTTTAAAGTACTGAATCATGTG 3240
3241 TGAATGAGACAGAGATGGCAAAATGAATGAGCACAACCAATTTCTTCTTCTTCTTCTTCTT 3300
3241 TGAATGAGACAGAGATGGCAAAATGAATGAGCACAACCAATTTCTTCTTCTTCTTCTTCTT 3300
3301 CAGATGACCTGATCTGCATCAGAGTGTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3360
3301 CAGATGACCTGATCTGCATCAGAGTGTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3360
3361 GGTGCTTCAAGAGCTGTGTGAGTGTGTGATCTCTGCCATCTCTGGCTGAGATATGC 3420
3361 GGTGCTTCAAGAGCTGTGTGAGTGTGTGATCTCTGCCATCTCTGGCTGAGATATGC 3420
3421 TGTCTGTGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3480
3421 TGTCTGTGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3480
3481 TAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3540
3481 TAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3540

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/01101
;; FILING DATE: FILED HEREWITH
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/005,156
;; FILING DATE: 15-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: GATES, EDWARD R.
;; REGISTRATION NUMBER: 31,616
;; REFERENCE/DOCKET NUMBER: B0819/7000WO
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/720-3500
;; TELEFAX: 617/720-2441
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1854 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: HOMO SAPIENS
;; CELL TYPE: LYMPHOMA
PCT-US94-01101-1

Query Match 16.5%; Score 774; DB 5; Length 1854;
Best Local Similarity 100.0%; Pred. No. 1.7e-289;
Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 703 GGCACACCGCTCCCTCAAGAAAGAGTGTGAGCGGGCCGGGAGGAGGAGGAGG 762
DB 1672 GGCACACCGCTCCCTCAAGAAAGAGTGTGAGCGGGCCGGGAGGAGGAGGAGG 1613
QY 763 GCAAGGAGGTGACCTCGACATTCAGGTGACATGATGACATTCAGTGCAGATGGAAC 822
DB 1612 GCAAGGAGGTGACCTCGACATTCAGGTGACATGATGACATTCAGTGCAGATGGAAC 1553
QY 823 AGCACATGAGCGCACTCCCAAGCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 882
DB 1552 AGCACATGAGCGCACTCCCAAGCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1493
QY 883 AGAAGCTGATGAGCAGTATGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 942
DB 1492 AGAAGCTGATGAGCAGTATGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1433
QY 943 AGACCTACACAGCAGCTGGTGGATGCCAGCTCCAGCAGGAGGAGGAGGAGGAGGAGG 1002
DB 1432 AGACCTACACAGCAGCTGGTGGATGCCAGCTCCAGCAGGAGGAGGAGGAGGAGGAGG 1373
QY 1003 AGGACAGAGGCGGACCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1062
DB 1372 AGGACAGAGGCGGACCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1313
QY 1063 AGAGGATGTGAGCTGTGATGAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1122
DB 1312 AGAGGATGTGAGCTGTGATGAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1253
QY 1123 ACACAGAGAGTTTGGAGGTTCCAGAACACACTTTCCAAAAGCAGCAGGAGGATTCACCA 1182
DB 1252 ACACAGAGAGTTTGGAGGTTCCAGAACACACTTTCCAAAAGCAGCAGGAGGATTCACCA 1193
QY 1183 CATTCAAGCAGGAGATGGAAGATCACTAAGAGATCAAGAGCTGGAAGAGGAGGAGGAGG 1242
DB 1192 CATTCAAGCAGGAGATGGAAGATCACTAAGAGATCAAGAGCTGGAAGAGGAGGAGGAGG 1133
QY 1243 CCATGTACCGTCCCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1302
DB 1132 CCATGTACCGTCCCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1073
QY 1303 AAACAGTCCGGGATTAAGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1362

DB 1072 AAACAGTCCGGGATTAAGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1013
QY 1363 TGTGCGGGGCACTGCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1422
DB 1012 TGTGCGGGGCACTGCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 953
QY 1423 CTGCTGGCCAGGGCTCCCTCACTGACAGTGCCTCTGAGAGGAGGAGGAGGAGGAGGAGG 1476
DB 952 CTGCTGGCCAGGGCTCCCTCACTGACAGTGCCTCTGAGAGGAGGAGGAGGAGGAGGAGG 899

RESULT 7
US-09-833-381-1334
; Sequence 1334, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1334
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(492)
; OTHER INFORMATION: n = A, T, C or G
US-09-833-381-1334

Query Match 4.6%; Score 218; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 7.2e-75;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2834 GCCCATAGCTGTTCTCTAAGCAGGTAAAGAAATGAATGAATGAATGAATGAATGAATGAAT 2893
DB 201 GCCCATAGCTGTTCTCTAAGCAGGTAAAGAAATGAATGAATGAATGAATGAATGAATGAAT 260
QY 2894 ATCCATAAATGGCTGCAGACAGTGTGTTCTGTCCTGTTCTGTCCTGTTCTGTCCTGTTCTGTCCT 2953
DB 261 ATCCATAAATGGCTGCAGACAGTGTGTTCTGTCCTGTTCTGTCCTGTTCTGTCCTGTTCTGTCCT 320
QY 2954 TAATCTATGCTACTGTGTAGAGCCATTCATATGCTGAATGTTCTGCTGCTGCTGCTGCTGCTGCT 3013
DB 321 TAATCTATGCTACTGTGTAGAGCCATTCATATGCTGAATGTTCTGCTGCTGCTGCTGCTGCTGCT 380
QY 3014 GCCAGGATATTAGCCAGTGTGTTGTCAGCAGTGTGTTGTCAGCAGTGTGTTGTCAGCAGTGTGTT 3051
DB 381 GCCAGGATATTAGCCAGTGTGTTGTCAGCAGTGTGTTGTCAGCAGTGTGTTGTCAGCAGTGTGTT 418

RESULT 8
US-09-621-976-16512
; Sequence 16512, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16512
; LENGTH: 421
; TYPE: DNA

ORGANISM: Homo sapiens
US-09-621-976-16512
Query Match 3.7%; Score 173; DB 4; Length 421;
Best Local Similarity 99.7%; Pred. No. 1.7e-57;
Matches 293; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 4217 TTTATTTGTTGTTGTTTCCATGAGGTTATCGACCATGGCTGAGCTCAGGCACTTC 4276
DB 1 TTTATTTGTTGTTGTTTCCATGAGGTTATCGACCATGGCTGAGCTCAGGCACTTC 60
QY 4277 TGTAGGAGACTGTTATTTCTGTAAAGATGTTTATTTAAACCTCTCCACCCCATCACGGT 4336
DB 61 TGTAGGAGACTGTTATTTCTGTAAAGATGTTTATTTAAACCTCTCTCCACCCCATCACGGT 120
QY 4337 GGCCTTGAGGCTGACCCGAGGCCAGTGGAGCTGCTGTTTCCACGGGGAGGGCCAA 4396
DB 121 GGCCTTGAGGCTGACCCGGA-GCCAGTGGAGCTGCTGTTTCCACGGGGAGGGCCAA 179
QY 4397 GGCCTTGAGCTGATCTCCAGCTGCTGCCAGGCTTCCGCTTGCACAGGCACAG 4456
DB 180 GGCCTTGAGCTGATCTCCAGCTGCTGCCAGGCTTCCGCTTGCACAGGCACAG 239
QY 4457 GTGGTCAACCCAGGACAGCCAGGCACTGCTCTCTTCCCTTCTCGGGGAA 4510
DB 240 GTGGTCAACCCAGGACAGCCAGGCACTGCTCTCTTCCCTTCTCGGGGAA 293
RESULT 9
US-08-979-608A-14
; Sequence 14, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 61...1731
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-979-608A-14
Query Match 1.3%; Score 59; DB 4; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAAACAAGAGGT 1408
DB 1456 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAAACAAGAGGT 1514
RESULT 10
US-09-517-849-14
; Sequence 14, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 61...1731
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-517-849-14
Query Match 1.3%; Score 59; DB 4; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAAACAAGAGGT 1408
DB 1456 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAAACAAGAGGT 1514

```
RESULT 11
US-09-616-289-14
; Sequence 14, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 03/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-09-616-289-14
; Query Match
; Best Local Similarity 1.3%; Score 59; DB 4; Length 4722;
; Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1350 CGGCTGAGAGGCTGTGCGGGCACTGCACAGAGGCGCAATGACCTGAACAAGAGGGT 1408
DB 1456 CGGCTGAGAGGCTGTGCGGGCACTGCACAGAGGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 12
US-08-859-998-309/c
; Sequence 309, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jokhadze, George
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 310:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 310:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
```

```
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 309:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-08-859-998-309
; Query Match
; Best Local Similarity 0.6%; Score 30; DB 2; Length 30;
; Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1165 GCAGCGAGGTATTCCACCATTCAGCAGG 1194
DB 30 GCAGCGAGGTATTCCACCATTCAGCAGG 1

RESULT 13
US-08-859-998-310
; Sequence 310, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jokhadze, George
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 310:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
```


FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-310
Query Match 0.6%; Score 30; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 814 AGATGGAACAGCACAAATGAGCGCAACTCCA 843
DB 1 AGATGGAACAGCACAAATGAGCGCAACTCCA 30
RESULT 14
US-09-225-928-309/c
Sequence 309, Application US/09225928
Patent No. 6352829
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
Jokhadze, George
Bibilashvilli, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,928
FILING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,998
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 309:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
SEQUENCE DESCRIPTION: SEQ ID NO: 309:
US-09-225-928-309
Query Match 0.6%; Score 30; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1165 GCAGCGAGGTATTCACACATTCACGAGG 1194
DB 30 GCAGCGAGGTATTCACACATTCACGAGG 1
RESULT 15
US-09-225-928-310

Sequence 310, Application US/09225928
Patent No. 6352829
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
Jokhadze, George
Bibilashvilli, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,928
FILING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,998
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
SEQUENCE DESCRIPTION: SEQ ID NO: 310:
US-09-225-928-310
Query Match 0.6%; Score 30; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 814 AGATGGAACAGCACAAATGAGCGCAACTCCA 843
DB 1 AGATGGAACAGCACAAATGAGCGCAACTCCA 30
RESULT 16
US-09-225-201B-309/c
Sequence 309, Application US/09225201B
Patent No. 6489455
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
Jokhadze, George
Bibilashvilli, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US


```

; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1502..1521
; OTHER INFORMATION: 99-26781-25.mis2, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1477..1497
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1905..1925
; OTHER INFORMATION: downstream amplification primer, complement
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1489..1513
; OTHER INFORMATION: 99-26781-25 probe
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 21,274..275
; OTHER INFORMATION: n=a, g, c or t
; US-09-539-333D-211

```

Query Match 0.6%; Score 28; DB 4; Length 3001;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

```

RESULT 19
US-09-508-824-17/c
; Sequence 17, Application US/09508824
; Patent No. 6635811
; GENERAL INFORMATION:
; APPLICANT: Flinham, John E
; APPLICANT: Gale, Michael D
; APPLICANT: Holdsworth, Michael J
; TITLE OF INVENTION: Pre-harvest Spr
; FILE REFERENCE: Newburt
; CURRENT APPLICATION NUMBER: US/09/5
; CURRENT FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: PCT/GB98/
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: GB 972006
; PRIOR FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 17
; LENGTH: 953
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-508-824-17

```

Query Match 0.6%; Score 27; DB 4; Length 953;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 27: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2381 TTTTTTTTTTTCTTTTTTTTTT 2407
pb 953 TTTTTTTTTTTCTTTTTTTTTT 927

RESULT 20
US-09-564-805-28
Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Sinaid, Jacques
; APPLICANT: Rommens, Johanna M.

```

; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc.feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22873-22917
; NAME/KEY: misc.feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 71631, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
; US-09-564-805-28

```

```
Query Match          0.6%; Score 27; DB 4; Length 26664;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 27: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

```

RESULT 21
US-09-548-797B-7/c
; Sequence 7, Application US/09548797B
; Patent No. 6683165
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES AND
; FILE REFERENCE: 2976-4039
; CURRENT APPLICATION NUMBER: US/09/548,797B
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 60/129,391
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 17000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-797B-7

```

1-08-731-272A-26
 Sequence 26, Application US/08731272A
 Patent No. 5837463
 GENERAL INFORMATION:
 APPLICANT: Tanaka, Torahiko
 APPLICANT: Katch, No. 5837483uyuki
 APPLICANT: Shimotohno, Kunitada
 TITLE OF INVENTION: NUCLEIC ACID OF C TYPE HEPATITIS VIRUS
 TITLE OF INVENTION: DERIVATION AND PROCESS FOR DETECTION VIRUS USING SAID
 TITLE OF INVENTION: NUCLEIC ACID
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/731,272A
 FILING DATE: 11-OCT-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP HEI. 7/268700
 FILING DATE: 17-OCT-1995

OY 4673 TTGCCTTAAAAA 4697
 DB 4132 TTGCCTTAAAAA 4156

RESULT 25

US-09-030-970-1
 ; Sequence 1, Application US/09030970
 ; Patent No. 6143519
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, ET AL.
 ; TITLE OF INVENTION: Human Endothelin Receptor
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/030,970
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/465,687
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MULLINS, J.G.
 ; REGISTRATION NUMBER: 33,073
 ; REFERENCE/DOCKET NUMBER: 325800-322 (PF137)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4156 BASE PAIRS
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: SINGLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: CDNA
 ; US-09-030-970-1

Query Match 0.5%; Score 25; DB 3; Length 4156;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4673 TTGCCTTAAAAA 4697
 DB 4132 TTGCCTTAAAAA 4156

RESULT 26

US-520-210-1
 ; Sequence 1, Application US/09520210
 ; Patent No. 6518404
 ; GENERAL INFORMATION:
 ; APPLICANT: Li et al.
 ; TITLE OF INVENTION: Human Endothelin-Bombesin Receptor
 ; FILE REFERENCE: PF137D2
 ; CURRENT APPLICATION NUMBER: US/09/520,210
 ; CURRENT FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: 09/030,970
 ; PRIOR FILING DATE: 1998-02-26
 ; PRIOR APPLICATION NUMBER: 08/465,687
 ; PRIOR FILING DATE: 1995-06-06
 ; PRIOR APPLICATION NUMBER: PCT/US94/11843

; PRIOR FILING DATE: 1994-10-17
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 4156
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1779)...(3620)
 ; OTHER INFORMATION:
 ; US-09-520-210-1

Query Match 0.5%; Score 25; DB 4; Length 4156;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4673 TTGCCTTAAAAA 4697
 DB 4132 TTGCCTTAAAAA 4156

RESULT 27

PCT-US94-11843-1
 ; Sequence 1, Application PC/TUS9411843
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, ET AL.
 ; TITLE OF INVENTION: Human Endothelin Receptor
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/11843
 ; FILING DATE: Concurrently
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FERRARO, GREGORY D.
 ; REGISTRATION NUMBER: 36,134
 ; REFERENCE/DOCKET NUMBER: 325800-192
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4156 BASE PAIRS
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: SINGLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: CDNA
 ; PCT-US94-11843-1

Query Match 0.5%; Score 25; DB 5; Length 4156;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4673 TTGCCTTAAAAA 4697
 DB 4132 TTGCCTTAAAAA 4156


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,975B
; FILING DATE: 10-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20894/80
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4q35
; US-08-370-975B-1
;
; Query Match 0.5%; Score 25; DB 1; Length 26764;
; Best Local Similarity 100.0%; Pred No. 1.8;
; Matches 25; Conservative 0; Mismatches 0; Indels
;
; QY 2383 TTTTTCCTTTTTCTTTTTTTTTT 2407
;      |||||
; DB 17485 TTTTTCCTTTTTCTTTTTTTTTT 17461
;
; RESULT 36
; US-08-781-891-207
; Sequence 207, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29604 base pairs
; TYPE: nucleic acid

```



```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-207
Query Match 0.5%; Score 25; DB 3; Length 29604;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2383 TTTTTCCTTTTCTTTTCTTTTCTTTT 2407
DB 17175 TTTTTCCTTTTCTTTTCTTTTCTTTT 17199

RESULT 37
US-09-618-166-207
; Sequence 207, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/618,166
; FILING DATE: 17-Jul-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McWaters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 240052.419C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4300
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-09-618-166-207
Query Match 0.5%; Score 25; DB 4; Length 29604;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2383 TTTTTCCTTTTCTTTTCTTTTCTTTT 2407
DB 17175 TTTTTCCTTTTCTTTTCTTTTCTTTT 17199

RESULT 38
US-09-040-984-26
; Sequence 26, Application US/09040984
; Patent No. 6210883
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
; OF LUNG CANCER
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,984
; FILING DATE: 18-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-282-6031
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-040-984-26
Query Match 0.5%; Score 24; DB 3; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAAATAAAAAA 4697
DB 410 TGCCTTAAAAAATAAAAAA 433

RESULT 39
US-09-123-912-26
; Sequence 26, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, TongTong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 26
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (395)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-26
Query Match 0.5%; Score 24; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 4674 TGCCTTAAAAA 4697
 Db 410 TGCCTTAAAAA 433

RESULT 40

US-09-643-597-26
 ; Sequence 26, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; CURRENT FILING DATE: 2000-08-21
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 434
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(434)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-643-597-26

Query Match 0.5%; Score 24; DB 4; Length 434;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAA 4697
 Db 410 TGCCTTAAAAA 433

RESULT 41

US-09-480-884A-26
 ; Sequence 26, Application US/09480884A
 ; Patent No. 6482597
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 434
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(434)
 ; OTHER INFORMATION: n = A,T,C or G

US-09-480-884A-26

Query Match 0.5%; Score 24; DB 4; Length 434;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAA 4697
 Db 410 TGCCTTAAAAA 433

RESULT 42

US-09-542-615A-26
 ; Sequence 26, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C8
 ; CURRENT APPLICATION NUMBER: US/09/542,615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 434
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(434)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-542-615A-26

Query Match 0.5%; Score 24; DB 4; Length 434;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAA 4697
 Db 410 TGCCTTAAAAA 433

RESULT 43

US-09-606-421B-26
 ; Sequence 26, Application US/09606421B
 ; Patent No. 6531315
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C9
 ; CURRENT APPLICATION NUMBER: US/09/606,421B
 ; CURRENT FILING DATE: 2000-06-28
 ; NUMBER OF SEQ ID NOS: 358
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 434
 ; TYPE: DNA
 ; ORGANISM: Homo sapien

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456
TELEPHONE: 206-622-4900
TELEFAX: 206-282-6031
TELEX:
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-040-984-74
Query Match 0.5%; Score 24; DB 3; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4674 TGCCTTAAAAA 4697
DB 410 TGCCTTAAAAA 433
RESULT 46
US-09-123-912-74
Sequence 74, Application US/09123912A
Patent No. 6312695
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C1
CURRENT APPLICATION NUMBER: US/09/123,912A
CURRENT FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: 09/040,802
PRIOR FILING DATE: 1998-03-18
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 74
LENGTH: 437
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (145)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (355)
OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-74
Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4674 TGCCTTAAAAA 4697
DB 410 TGCCTTAAAAA 433
RESULT 47
US-09-643-597-74
Sequence 74, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.

NAME/KEY: misc feature
LOCATION: (1)-(434)
OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-26
Query Match 0.5%; Score 24; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4674 TGCCTTAAAAA 4697
DB 410 TGCCTTAAAAA 433
RESULT 44
US-09-221-107-26
Sequence 26, Application US/09221107
Patent No. 6660838
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221,107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 26
LENGTH: 434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (395)
OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-26
Query Match 0.5%; Score 24; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4674 TGCCTTAAAAA 4697
DB 410 TGCCTTAAAAA 433
RESULT 45
US-09-040-984-74
Sequence 74, Application US/09040984
Patent No. 6210883
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF LUNG CANCER
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(437)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-74

```

```

Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred.No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4674 TGCCTTAAAAA 4697
Db 410 TGCCTTAAAAA 433

```

```

RESULT 48
US-09-480-884A-74
; Sequence 74, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(437)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-74

```

```

Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred.No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4674 TGCCTTAAAAA 4697
Db 410 TGCCTTAAAAA 433

```

```

RESULT 49
US-09-542-615A-74
; Sequence 74, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun

```

```

; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(437)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-74

```

```

Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred.No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4674 TGCCTTAAAAA 4697
Db 410 TGCCTTAAAAA 433

```

```

RESULT 50
US-09-606-421B-74
; Sequence 74, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(437)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-74

```

```

Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred.No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4674 TGCCTTAAAAA 4697
Db 410 TGCCTTAAAAA 433

```

```

RESULT 51
US-09-221-107-74
; Sequence 74, Application US/09221107

```

Patent No. 6660838
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221,107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 74
LENGTH: 437
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (145)
OTHER INFORMATION: Where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (355)
OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-74

Query Match 0.5%; Score 24; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAA 4697
DB 410 TGCCTTAAAAA 433

RESULT 52
US-07-885-970A-15
Sequence 15, Application US/07885970A
Patent No. 5495070
GENERAL INFORMATION:
APPLICANT: John, Maliyakal E.
TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nicholas J. Seay, Quarles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,970A
FILING DATE: 19920518
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/617,239
FILING DATE: 21-NOV-1990
PRIOR APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
TELEPHONE: (608) 283-2478
TELEFAX: (608) 251-5139
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1080 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum
STRAIN: Coker 312
DEVELOPMENTAL STAGE: 10 day old fiber cells
TISSUE TYPE: fiber cells
IMMEDIATE SOURCE:
LIBRARY: CKFB10
CLONE: B6
US-07-885-970A-15

Query Match 0.5%; Score 24; DB 1; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4673 TTGCTTAAAAA 4696
DB 1057 TTGCTTAAAAA 1080

RESULT 53
US-08-298-687A-15
Sequence 15, Application US/08298687A
Patent No. 5521078
GENERAL INFORMATION:
APPLICANT: John, Maliyakal E.
TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
TITLE OF INVENTION: PLANTS FOR ALTERED FIBER
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nicholas J. Seay, Quarles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/298,687A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/617,239
FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
TELEPHONE: (608) 283-2478
TELEFAX: (608) 251-5139
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1080 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum
STRAIN: Coker 312
DEVELOPMENTAL STAGE: 10 day old fiber cells

TISSUE TYPE: fiber cells
IMMEDIATE SOURCE:
LIBRARY: CKFB10
CLONE: B6
US-08-298-687A-15

Query Match 0.5%; Score 24; DB 1; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4673 TTGCCTTAAAAA 4696
DB 1057 TTGCCTTAAAAA 1080

RESULT 54

US-08-530-797-16
Sequence 16, Application US/08530797
Patent No. 559718
GENERAL INFORMATION:
APPLICANT: John, Maliyakal E.
APPLICANT: Umbeck, Paul F.
APPLICANT: Brill, Winston J.
TITLE OF INVENTION: GENETICALLY ENGINEERED COTTON PLANTS
TITLE OF INVENTION: FOR ALTERED FIBER
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles and Brady
STREET: P.O. BOX 2113
CITY: MADISON
STATE: WISCONSIN
COUNTRY: U.S.A.
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 800Kb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,797
FILING DATE: 20-SEP-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/617,239
FILING DATE: 21-NOV-90
APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-88
ATTORNEY/AGENT INFORMATION:
NAME: Nicholas J. Seay
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 1122990245
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum
STRAIN: Coker 312
DEVELOPMENTAL STAGE: 10 day old fiber cells
TISSUE TYPE: fiber cells
IMMEDIATE SOURCE:
LIBRARY: CKFB10
CLONE: B6
US-08-530-797-16

Query Match 0.5%; Score 24; DB 1; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4673 TTGCCTTAAAAA 4696
DB 1057 TTGCCTTAAAAA 1080

RESULT 55

US-08-298-829-15
Sequence 15, Application US/08298829
Patent No. 5620882
GENERAL INFORMATION:
APPLICANT: John, Maliyakal E.
TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
TITLE OF INVENTION: PLANTS FOR ALTERED FIBER
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nicholas J. Seay, Quarles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/298,829
FILING DATE: 19-OCT-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/885,970
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/617,239
FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 283-2478
TELEFAX: (608) 251-5139
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1080 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum
STRAIN: Coker 312
DEVELOPMENTAL STAGE: 10 day old fiber cells
TISSUE TYPE: fiber cells
IMMEDIATE SOURCE:
LIBRARY: CKFB10
CLONE: B6
US-08-298-829-15

Query Match 0.5%; Score 24; DB 1; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4673 TTGCCTTAAAAA 4696
DB 1057 TTGCCTTAAAAA 1080

```

; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 60
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-60

Query Match 0.5%; Score 24; DB 4; Length 1272;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2384 TTTTTCCTTTTCTTTTCTTTT 2407
Db 1256 TTTTTCCTTTTCTTTTCTTTT 1233

RESULT 58
US-08-395-800A-7/c
; Sequence 7, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES,DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; TITLE OF INVENTION: GENOTYPING A PERSON
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248955 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..1092
; US-08-395-800A-7

```

Query Match 0.5%; Score 24; DB 1; Length 2115;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2367 ATTCTTTTCTTTTCTTTTCTTTTCTTTT 2390
|||||
Db 1552 ATTCTTTTCTTTTCTTTTCTTTTCTTTT 1529

RESULT 59
US-09-684-708A-4
; Sequence 4, Application US/09684708A
; Patent No. 6602851
; GENERAL INFORMATION:
; APPLICANT: Carroil, Steven, M.D., Ph.D
; TITLE OF INVENTION: SMDF Neuregulin Splice Variant Isoforms and Uses
; FILE REFERENCE: D6240
; CURRENT APPLICATION NUMBER: US/09/684,708A
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/158,622
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 4
; LENGTH: 2540
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 370..2458
; OTHER INFORMATION: SMDF_2a amino acid sequence
US-09-684-708A-4

Query Match 0.5%; Score 24; DB 4; Length 2540;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2379 TTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2402
|||||
Db 175 TTTTCTTTTCTTTTCTTTTCTTTTCTTTT 198

RESULT 60
US-09-539-333D-195
; Sequence 195, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 195
; LENGTH: 3001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1501
; OTHER INFORMATION: 99-26225-148 : polymorphic base G or T
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1481..1500
; OTHER INFORMATION: 99-26225-148.misl,
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1502..1520
; OTHER INFORMATION: 99-26225-148.mis2, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1355..1375
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1805..1825
; OTHER INFORMATION: downstream amplification primer, complement
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1489..1513
; OTHER INFORMATION: 99-26225-148 probe
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 37,514,2455
; OTHER INFORMATION: n=a, g, c or t
US-09-539-333D-195

Query Match 0.5%; Score 24; DB 4; Length 3001;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2367 ATTCTTTTCTTTTCTTTTCTTTTCTTTT 2390
|||||
Db 2427 ATTCTTTTCTTTTCTTTTCTTTTCTTTT 2450

RESULT 61
US-08-420-306D-1/c
; Sequence 1, Application US/08420306D
; Patent No. 6605274
; GENERAL INFORMATION:
; APPLICANT: DILLMANN, WOLFGANG H.
; APPLICANT: GIORDANO, FRANK
; APPLICANT: MESTRIL, RUBEN
; APPLICANT: HAMMOND, KIRK
; TITLE OF INVENTION: METHOD FOR IN VIVO REGULATION OF CARDIAC MUSCLE
; FILE REFERENCE: 041673-2031
; CURRENT APPLICATION NUMBER: US/08/420,306D
; CURRENT FILING DATE: 1995-04-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 1
; LENGTH: 3517
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (995)..(1294)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base

LOCATION: (1320)...(1322)
OTHER INFORMATION: a, t, c, g, other or unknown
FEATURE:
NAME/KEY: modified base
LOCATION: (1541)...(1542)
OTHER INFORMATION: a, t, c, g, other or unknown
US-08-420-306D-1

Query Match 0.5%; Score 24; DB 4; Length 3517;
Best Local Similarity 100.0%; Pred. No. 4.9; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 4674 TGCCTTAAAAA...AAAAA 4697
Db 439 TGCCTTAAAAA...AAAAA 416

RESULT 62
US-09-254-325-1/c
Sequence 1, Application US/09254325
Patent No. 6030607
GENERAL INFORMATION:
APPLICANT: Veenstra, Annemarie E.
TITLE OF INVENTION: ENHANCED EXPRESSION OF
TITLE OF INVENTION: PROTEOLYTIC ENZYMES IN KOJI MOLDS
NUMBER OF SEQUENCES: 17
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION NUMBER: US/09/254,325
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4657 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: exon
LOCATION: 1189..1604
FEATURE:
NAME/KEY: intron
LOCATION: 1605..1703
FEATURE:
NAME/KEY: exon
LOCATION: 1704..3846
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1189..3480
OTHER INFORMATION: /label= TRUNCATED-AREA
OTHER INFORMATION: /note = "AREA IS TRUNCATED IMMEDIATELY
OTHER INFORMATION: DOWNSTREAM THE SEQUENCE ENCODING
OTHER INFORMATION: A DNA BINDING DOMAIN"

US-09-254-325-1
Query Match 0.5%; Score 24; DB 3; Length 4657;
Best Local Similarity 100.0%; Pred. No. 4.8; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 2367 ATTCTTTTCTTTT...TTTTT 2390
Db 478 ATTCTTTTCTTTT...TTTTT 455

RESULT 63
US-08-222-617A-1
Sequence 1, Application US/08222617A
Patent No. 582879
GENERAL INFORMATION:

APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Doehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Eduard P.
APPLICANT: Montenegro, Eduardo P.
TITLE OF INVENTION: A Method for Influencing Beta-Lactam
TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
TITLE OF INVENTION: Quantities of ACV Synthetase
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/222,617A
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 12364 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Penicillium chrysogenum
FEATURE:
NAME/KEY: CDS
LOCATION: 264..11600
OTHER INFORMATION: /function= "enzyme"
OTHER INFORMATION: /product= "ACV Synthetase"
US-08-222-617A-1

Query Match 0.5%; Score 24; DB 2; Length 12364;
Best Local Similarity 100.0%; Pred. No. 4.5; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 2384 TTTTTCCTTTTCTTTT...TTTTT 2407
Db 11997 TTTTTCCTTTTCTTTT...TTTTT 12020

RESULT 64
US-09-922-445-1/c
Sequence 1, Application US/09922445
Patent No. 6528268
GENERAL INFORMATION:
APPLICANT: Andersson, Maria K.
APPLICANT: Berglund, Lars G. T.
APPLICANT: Reneland, Rikard H.
APPLICANT: Adam, Gail I. R.
TITLE OF INVENTION: REAGENTS AND METHODS FOR DETECTION OF HEART FAILURE
FILE REFERENCE: GG126US
CURRENT APPLICATION NUMBER: US/09/922,445
CURRENT FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin version 3.1

```

; SEQ ID NO 1
; LENGTH: 38653
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(26156)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (24801)..(24801)
; OTHER INFORMATION: nucleotide 24801 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or G
; NAME/KEY: misc feature
; LOCATION: (24941)..(24941)
; OTHER INFORMATION: nucleotide 24941 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: T or C
; NAME/KEY: exon
; LOCATION: (26157)..(26252)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26253)..(26401)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (26402)..(26543)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26544)..(27024)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (27025)..(27178)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (27179)..(30519)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (27645)..(27645)
; OTHER INFORMATION: nucleotide 27645 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: C or G
; NAME/KEY: exon
; LOCATION: (30520)..(30681)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (30682)..(30894)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (30895)..(31027)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31028)..(31747)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (31748)..(31841)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31842)..(32400)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (32163)..(32163)
; OTHER INFORMATION: nucleotide 32163 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or C
; NAME/KEY: exon
; LOCATION: (32401)..(32528)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (32529)..(33414)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (32614)..(32614)
; OTHER INFORMATION: nucleotide 32614 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or G
; NAME/KEY: exon
; LOCATION: (33415)..(33597)
; OTHER INFORMATION:

```

```

; NAME/KEY: Intron
; LOCATION: (33598)..(34314)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (34315)..(34588)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (34589)..(36404)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (36405)..(36523)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (36524)..(38341)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (38342)..(38653)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank/AC004923
; DATABASE ENTRY DATE: 1999-12-21
; RELEVANT RESIDUES: (1)..(38653)
US-09-922-445-1

Query Match          0.5%; Score 24; DB 4; Length 38653;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4674 TGCCTTAAAAA 4697
      |||||
Db 1532 TGCCTTAAAAA 1509

RESULT 65
US-09-146-053-5/c
; Sequence 5, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkie, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146.053A
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 44453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-5

Query Match          0.5%; Score 24; DB 4; Length 44453;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2373 TTTTCTTTT 2396
      |||||
Db 18568 TTTTCTTTT 18565

RESULT 66
US-09-798-096-10/c
; Sequence 10, Application US/09798096
; Patent No. 6399378
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL2 EXPRESSION
; FILE REFERENCE: RTS-0207

```

; OTHER INFORMATION: RH830

; SOFTWARE: FastSEQ for Windows Version 1.0.0

```

; TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE:  210121.478C12
; CURRENT APPLICATION NUMBER:  US/09/671,325
; CURRENT FILING DATE:  2000-09-26
; NUMBER OF SEQ ID NOS:  1825
; SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 612
; LENGTH:  225
; TYPE:  DNA
; ORGANISM:  Homo sapien
; FEATURE:
; NAME/KEY:  misc feature
; LOCATION:  (1)..(225)
; OTHER INFORMATION:  n = A,T,C or G
; US-09-671-325-612

Query Match      0.5%; Score 23; DB 4; Length 225;
Best Local Similarity 100.0%; Pred.No.14;
Matches 23; Conservative 0; Mismatches 0; Indels 0;

QY      2387 TTTTCTTTTCTTTTCTTTTTCG 2409
          |||||
DB      215 TTTTCTTTTCTTTTCTTTTTCG 193

RESULT 73
US-09-589-184-612/C
; Sequence 612, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT:  Wang, Tongtong
; APPLICANT:  Bangur, Chaitanya S.
; APPLICANT:  Lodes, Michael A.
; APPLICANT:  Fanger, Gary
; APPLICANT:  Vedwick, Tom
; APPLICANT:  Carter, Darrick
; APPLICANT:  Retter, Marc
; APPLICANT:  Mannion, Jane
; TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AN
; FILE REFERENCE:  210121.478C8
; CURRENT APPLICATION NUMBER:  US/09/589,184
; CURRENT FILING DATE:  2000-06-05
; NUMBER OF SEQ ID NOS:  827
; SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 612
; LENGTH:  225
; TYPE:  DNA
; ORGANISM:  Homo sapien
; FEATURE:
; NAME/KEY:  misc feature
; LOCATION:  (1)..(225)
; OTHER INFORMATION:  n = A,T,C or G
; US-09-589-184-612

Query Match      0.5%; Score 23; DB 4; Length 225;
Best Local Similarity 100.0%; Pred.No.14;
Matches 23; Conservative 0; Mismatches 0; Indels 0;

QY      2387 TTTTCTTTTCTTTTCTTTTTCG 2409
          |||||
DB      215 TTTTCTTTTCTTTTCTTTTTCG 193

RESULT 74
US-08-520-678A-29
; Sequence 29, Application US/08520678A
; Patent No. 5874565
; GENERAL INFORMATION:
; APPLICANT:  Rice, Charles M.
; APPLICANT:  Kolyhalov, Alexander A.
; TITLE OF INVENTION:  NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
; TITLE OF INVENTION:  C VIRUS GENOME AND DIAGNOSTIC AND THER

```

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howell & Haferkamp, L.C.
STREET: 7733 Forsyth Blvd., Suite 1400
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/520,678A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Henderson, Melodie W.
REGISTRATION NUMBER: 37,848
REFERENCE/DOCKET NUMBER: 6029-6836
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-520-678A-29

Query Match 0.5%; Score 23; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCCTTTTCCTTT 2400
DB 174 TTTTTCCTTTTCCTTTTCCTTT 196

RESULT 75
US-08-897-126-29
Sequence 29, Application US/08897126
Patent No. 6297003
GENERAL INFORMATION:
APPLICANT: Rice, Charles M.
TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howell & Haferkamp, L.C.
STREET: 7733 Forsyth Blvd., Suite 1400
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/520,678
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Henderson, Melodie W.
REGISTRATION NUMBER: 37,848
REFERENCE/DOCKET NUMBER: 6029-6836
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-897-126-29

Query Match 0.5%; Score 23; DB 3; Length 260;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCCTTTTCCTTT 2400
DB 174 TTTTTCCTTTTCCTTTTCCTTT 196

RESULT 76
US-08-687-080-85
Sequence 85, Application US/08687080
Patent No. 5965427
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 5' END OF INTRON 13 OF RAD50 GENOMIC
INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-85

Query Match 0.5%; Score 23; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2384 TTTTCTTTCTTTCTTTCTTTCTTTCTTT 2406
DB 180 TTTTCTTTCTTTCTTTCTTTCTTTCTTT 202

RESULT 77

US-09-328-111-190
Sequence 1818, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054P2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 1818
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 62..220
NAME/KEY: sig peptide
LOCATION: 62..121
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 4.40000009536743
OTHER INFORMATION: seq TFFLISETIA/DL
US-09-621-976-1818

Query Match 0.5%; Score 23; DB 4; Length 519;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2385 TTTTCTTTCTTTCTTTCTTTCTTTCTTT 2407
DB 202 TTTTCTTTCTTTCTTTCTTTCTTTCTTT 180

RESULT 78

US-09-328-111-190
Sequence 190, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Ascle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 190

LENGTH: 699
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(699)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-190

Query Match 0.5%; Score 23; DB 3; Length 699;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2374 TTTCTTTTCTTTCTTTCTTTCTTTCTTT 2396
DB 31 TTTCTTTTCTTTCTTTCTTTCTTTCTTT 53

RESULT 79

US-09-257-179-17
Sequence 17, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P2015P1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1993-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17
LENGTH: 925
TYPE: DNA
ORGANISM: Homo sapiens
US-09-257-179-17

Query Match 0.5%; Score 23; DB 4; Length 925;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4675 GCCTTAAAAAATAAAAAAAAAAAAAA 4697
DB 898 GCCTTAAAAAATAAAAAAAAAAAAAA 920

RESULT 80

US-08-047-041A-24
Sequence 24, Application US/08047041A
Patent No. 5527676
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Baker, Suzanne J.
APPLICANT: Fearon, Eric R.
APPLICANT: Nigro, Janice M.
TITLE OF INVENTION: Detection of Loss of the Wild-Type p53
FILE REFERENCE: Gene
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20001.4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/047,041A
FILING DATE: 22-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US 07/928,661
APPLICATION NUMBER: US 07/928,661
FILING DATE: 17-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/446,584
FILING DATE: 06-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,566
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42917
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1288 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: exon 11
PUBLICATION INFORMATION:
AUTHORS: Lamb,
JOURNAL: Mol. Cell. Biol.
VOLUME: 6
ISSUE: 5
PAGES: 1379-1385
DATE: 1986
US-08-047-041A-24

Query Match 0.5%; Score 23; DB 1; Length 1288;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2378 TTTTTCCTTTTCTTTTCTTTTCTTT 2400
Db 841 TTTTTCCTTTTCTTTTCTTTTCTTT 863

RESULT 81
US-09-417-278A-1
Sequence 1, Application US/09417278A
Patent No. 6429298
GENERAL INFORMATION:
APPLICANT: Ellington, Andrew D.
APPLICANT: Matsumura, Ichiro
TITLE OF INVENTION: ASSAYS FOR IDENTIFYING FUNCTIONAL ALTERATIONS IN THE
FILE REFERENCE: INRP:070
CURRENT APPLICATION NUMBER: US/09/417,278A
CURRENT FILING DATE: 1999-10-13
PRIOR APPLICATION NUMBER: PCT/US99/24072
PRIOR FILING DATE: 1999-10-13
PRIOR APPLICATION NUMBER: 60/103,930
PRIOR FILING DATE: 1998-10-01

NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
US-09-417-278A-1

Query Match 0.5%; Score 23; DB 4; Length 1288;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2378 TTTTTCCTTTTCTTTTCTTTTCTTT 2400
Db 841 TTTTTCCTTTTCTTTTCTTTTCTTT 863

RESULT 82
US-08-047-041A-11
Sequence 11, Application US/08047041A
Patent No. 5527676
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Baker, Suzanne J.
APPLICANT: Fearon, Eric R.
APPLICANT: Nigro, Janice M.
TITLE OF INVENTION: Detection of Loss of the Wild-Type p53
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001.4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/047,041A
FILING DATE: 22-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,661
FILING DATE: 17-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/446,584
FILING DATE: 06-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,566
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42917
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:

CHROMOSOME/SEGMENT: exon 11
PUBLICATION INFORMATION:
AUTHORS: Buchman, et al.
JOURNAL: Gene
VOLUME: 70
PAGES: 245-252
DATE: 1988
US-08-047-041A-11

Query Match 0.5%; Score 23; DB 1; Length 1316;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCTTTTCTTTT 2400
DB 863 TTTTTCCTTTTCTTTTCTTTT 885

RESULT 83

US-08-795-006A-31
Sequence 31, Application US/08795006A
Patent No. 5840579
GENERAL INFORMATION:
APPLICANT: Boeke, Jef
APPLICANT: Brachmann, Rainer
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53
TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08795,006A
FILING DATE: 05-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.03170
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-795-006A-31

Query Match 0.5%; Score 23; DB 2; Length 1316;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCTTTTCTTTT 2400
DB 863 TTTTTCCTTTTCTTTTCTTTT 885

RESULT 84

US-08-795-006A-31
Sequence 31, Application US/08795006A
Patent No. 5840579
GENERAL INFORMATION:
APPLICANT: Boeke, Jef
APPLICANT: Brachmann, Rainer
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53
TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08795,006A
FILING DATE: 05-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.03170
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-795-006A-31

RESULT 85

US-08-413-135-3/c
Sequence 3, Application US/08413135
Patent No. 5689042
GENERAL INFORMATION:
APPLICANT: Amasino, Richard M
APPLICANT: Gan, Sushang
TITLE OF INVENTION: Transgenic Plants with Altered
TITLE OF INVENTION: Senescence Characteristics
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

US-09-184-073-31
Sequence 31, Application US/09184073
Patent No. 6183964
GENERAL INFORMATION:
APPLICANT: Boeke, Jef
APPLICANT: Brachmann, Rainer
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53
TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09184,073
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,006
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.03170
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-184-073-31

Query Match 0.5%; Score 23; DB 3; Length 1316;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCTTTTCTTTT 2400
DB 863 TTTTTCCTTTTCTTTTCTTTT 885

RESULT 85
US-08-413-135-3/c
Sequence 3, Application US/08413135
Patent No. 5689042
GENERAL INFORMATION:
APPLICANT: Amasino, Richard M
APPLICANT: Gan, Sushang
TITLE OF INVENTION: Transgenic Plants with Altered
TITLE OF INVENTION: Senescence Characteristics
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible


```
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/413,135
;; FILING DATE:
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Seay, Nicholas J.
;; REGISTRATION NUMBER: 27,386
;; REFERENCE/DOCKET NUMBER: 960296.94908
;; TELEPHONE: 608-251-5000
;; TELEFAX: 608-251-9166
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1974 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "SAG13 Promoter DNA"
US-08-413-135-3

Query Match 0.5%; Score 23; DB 1; Length 1974;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTCTTTTCTTTTCTTTTCTTTT 2390
Db 1246 TTCTTTTCTTTTCTTTTCTTTT 1224

RESULT 86
US-08-971-395-3/c
Sequence 3, Application US/08971395
Patent No. 6359197
GENERAL INFORMATION:
APPLICANT: Amasino, Richard M
APPLICANT: Gan, Susheng
TITLE OF INVENTION: Transgenic Plants with Altered
TITLE OF INVENTION: Senescence Characteristics
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,395
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.94908
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1974 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
```

```
US-08-971-395-3
Query Match 0.5%; Score 23; DB 4; Length 1974;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTCTTTTCTTTTCTTTTCTTTT 2390
Db 1246 TTCTTTTCTTTTCTTTTCTTTT 1224

RESULT 87
US-09-277-196-3
Sequence 3, Application US/09277196
Patent No. 6476206
GENERAL INFORMATION:
APPLICANT: Irink, Barry
APPLICANT: Jen, Jin
APPLICANT: Ratovitski, Edward
APPLICANT: Sidransky, David
TITLE OF INVENTION: p40 Protein Acts as an Oncogene
FILE REFERENCE: 01107.79765
CURRENT APPLICATION NUMBER: US/09/277,196
CURRENT FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: 60/079736
EARLIER FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 2065
TYPE: DNA
ORGANISM: Homo sapiens
US-09-277-196-3

Query Match 0.5%; Score 23; DB 4; Length 2065;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTCTTTTCTTTTCTTTTCTTT 2400
Db 1643 TTTTCTTTTCTTTTCTTTTCTTT 1665

RESULT 88
US-09-023-655-1359
Sequence 1359, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Sealhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1359:
SEQUENCE CHARACTERISTICS:
LENGTH: 2066 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G35215
US-09-023-655-1359

Query Match 0.5%; Score 23; DB 4; Length 2066;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2378 TTTTTCCTTTTCCTTTTCCTTTT 2400
|||||
Db 1641 TTTTTCCTTTTCCTTTTCCTTTT 1663

RESULT 89
US-08-539-304A-5
Sequence 5, Application US/08539304A
Patent No. 5792933
GENERAL INFORMATION:
APPLICANT: MA, DIN-POW
TITLE OF INVENTION: FIBER-SPECIFIC PROTEIN EXPRESSION IN THE
MEDIUM TYPE: COTTON PLANT
OPERATING SYSTEM: 7
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 JEFFERSON DAVIS HWY. SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/539,304A
FILING DATE: 04-OCT-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN, OBLON F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2343-037-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(741..1093, 1220..1226)
NAME/KEY: intron
LOCATION: 1094..1219

US-08-539-304A-5

Query Match 0.5%; Score 23; DB 1; Length 2646;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTTCTTTTCCTTTTCCTTTTCCTTTT 2390
|||||
Db 1465 TTTCTTTTCCTTTTCCTTTTCCTTTT 1487

RESULT 90
US-09-018-628-17
Sequence 17, Application US/09018628
Patent No. 5917019
GENERAL INFORMATION:
APPLICANT: de Lange, Titia
APPLICANT: van Steensel, Bas
APPLICANT: Bianchi, Alessandro
TITLE OF INVENTION: AN ALTERED TELOMERE REPEAT BINDING
MEDIUM TYPE: FACTOR AND THERAPEUTIC USE THEREOF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/018,628
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-191 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2907 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-09-018-628-17

Query Match 0.5%; Score 23; DB 2; Length 2907;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTTCTTTTCCTTTTCCTTTTCCTTTT 2390
|||||
Db 1700 TTTCTTTTCCTTTTCCTTTTCCTTTT 1722

RESULT 91
US-09-273-378-17
Sequence 17, Application US/09273378
Patent No. 6020166
GENERAL INFORMATION:
APPLICANT: de Lange, Titia
APPLICANT: van Steensel, Bas
APPLICANT: Bianchi, Alessandro

```
Query Match          0.5%; Score 23; DB 3; Length 2907;
Best Local Similarity 100.0%; Pred. NO.12;
Matches 23; Conservative 0; Mismatches 0; Indels

QY      2368 TTCTTTTTCCTTTTTTTTTTTT 2390
        |||||
Dd      1700 TTTCTTTTCCTTTTTTTTTTT 1722

RESULT 92
US-09-018-635-26
; Sequence 26, Application US/09018635
; Patent No 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

Query Match 0.5%; Score 23; DB 4; Length 2907;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels

QY 2368 TTTCTTTTCTTTTCTTTTCTTTTCTTTT 2390
 Db 1700 TTTCTTTTCTTTTCTTTTCTTTTCTTTT 1722

RESULT 94

US-09-912-962-26
 ; Sequence 26, Application US/09912962
 ; Patent No. 6586577
 ; GENERAL INFORMATION:
 ; APPLICANT: de Lange, Titia
 ; Broccoli, Dominique
 ; Smogorzewska, Agata
 ; TITLE OF INVENTION: TLOMERE REPEAT BINDING FACTOR AND
 ; DIAGNOSTIC AND THERAPEUTIC USE THEREOF
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: KLAUBER & JACKSON
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/912,962
 ; FILING DATE: 25-Jul-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/018,635
 ; FILING DATE: 04-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: David A. Jackson
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2907 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; DESCRIPTION: TRP2
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-912-962-26
 ; Sequence 26, Application US/09912962
 ; Patent No. 6586577
 ; GENERAL INFORMATION:
 ; APPLICANT: de Lange, Titia
 ; Broccoli, Dominique
 ; Smogorzewska, Agata
 ; TITLE OF INVENTION: TLOMERE REPEAT BINDING FACTOR AND
 ; DIAGNOSTIC AND THERAPEUTIC USE THEREOF
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: KLAUBER & JACKSON
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/912,962
 ; FILING DATE: 25-Jul-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/018,635
 ; FILING DATE: 04-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: David A. Jackson
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2907 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; DESCRIPTION: TRP2
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Query Match 0.5%; Score 23; DB 4; Length 2907;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTTCTTTTCTTTTCTTTTCTTTTCTTTT 2390
 Db 1700 TTTCTTTTCTTTTCTTTTCTTTTCTTTT 1722

RESULT 95

US-09-620-312D-554
 ; Sequence 554, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yuning
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; CURRENT FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt FL_genes Version 1.0
 ; SEQ ID NO 554
 ; LENGTH: 3845
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (408)..(2120)
 ; US-09-620-312D-554

Query Match 0.5%; Score 23; DB 4; Length 3845;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2385 TTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2407
 Db 3321 TTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3343

RESULT 96

US-09-620-312D-554/c
 ; Sequence 554, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yuning
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; CURRENT FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 12:36:00 ; Search time 11448.5 Seconds
(without alignments)
12251.597 Million cell updates/sec

Title: US-10-023-523-17

Perfect score: 4697

Sequence: 1 caaaagagcccgagacaa.....ttaaaaaaaaaaaaaaa 4697

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_esti:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	623	13.3	759	14	CF995807
2	572	12.2	1085	12	BM800456
3	556	11.8	835	14	CD655044
c 4	548	11.7	895	13	BU528395
					CF995807 AGENCOURT
					BM800456 AGENCOURT
					CD655044 AGENCOURT
					BU528395 AGENCOURT

5	538	11.5	606	12	BI006817
c 6	522	11.1	657	14	CA389551
c 7	489	10.4	591	14	CA417053
8	485	10.3	570	10	BF829485
9	472	10.0	799	13	BU849604
10	463	9.9	489	12	BM689900
11	463	9.9	597	14	CB384172
12	462	9.8	817	12	BG761234
c 13	451	9.6	713	13	BU732511
c 14	448	9.5	550	12	BM052772
c 15	445	9.5	598	13	BU729123
c 16	444	9.5	920	13	BQ96779
c 17	443	9.4	545	10	BF939517
18	443	9.4	880	13	BU506849
c 19	419	8.9	521	10	BE857092
20	419	8.9	932	12	BG827157
21	419	8.9	941	13	BQ328861
22	417	8.9	724	12	BI260369
23	410	8.7	890	13	BU194071
24	408	8.7	1061	13	BQ643746
25	406	8.6	773	10	BF692313
26	392	8.3	426	9	AV709674
27	387	8.2	543	14	CA313019
c 28	387	8.2	721	14	CA423847
c 29	386	8.2	488	9	AI498183
c 30	386	8.2	706	13	BU615776
31	386	8.2	970	13	EX370444
32	384	8.2	1089	12	BM452208
33	383	8.2	539	10	BF720458
34	383	8.2	1142	14	CD242362
35	381	8.1	1039	12	BM479937
c 36	380	8.1	482	9	AI342516
c 37	379	8.1	510	10	BF478077
38	379	8.1	954	10	BE876516
c 39	374	8.0	476	9	AI146364
40	374	8.0	875	13	BQ231672
41	373	7.9	410	10	BG056649
c 42	373	7.9	475	9	AI419038
43	373	7.9	892	13	BQ643720
44	370	7.9	583	13	BQ219653
45	365	7.8	772	12	BG400724
46	360	7.7	462	12	BM793609
c 47	359	7.6	487	10	AW504045
c 48	359	7.6	506	9	AI830542
49	358	7.6	790	14	CK000671
50	356	7.6	496	9	AI121132
51	355	7.6	1318	12	BM455349
c 52	354	7.5	456	9	AI394288
53	352	7.5	555	12	BM053066
c 54	343	7.3	445	10	BE858528
c 55	343	7.3	456	14	CB161278
56	343	7.3	729	12	BI868824
57	343	7.3	843	12	BI523918
58	343	7.3	891	10	BE892847
59	343	7.3	922	12	BG743552
60	343	7.3	927	12	BG169290
61	343	7.3	943	12	BM459633
62	343	7.3	988	12	BM471221
c 63	339	7.2	441	9	AI740467
64	339	7.2	878	13	BQ680816
65	335	7.1	541	14	CA389550
66	335	7.1	651	12	BM702342
c 67	334	7.1	849	13	EX331009
68	326	6.9	379	9	AI923296
69	323	6.9	568	12	BM833788
70	320	6.8	535	10	BE883980
c 71	319	6.8	472	9	AI139350
72	318	6.8	397	13	BU429029
73	316	6.7	524	14	CB130906
74	315	6.7	628	10	AW812918
75	314	6.7	807	14	CF786073
76	312	6.6	1333	12	BM924139
c 77	311	6.6	413	10	BF434394

78 306 6.5 1195 13 BX362957 BX362957
79 304 6.5 487 10 AW474425 XS95901.X
80 292 6.2 496 9 AI813503 WJ64902.X
81 291 6.2 457 9 AI335551 QT24H05.X
82 291 6.2 616 9 AA115345 T110F10.X
83 287 6.1 403 13 BU752065 UI-1-BB0-
84 287 6.1 432 12 BM766498 K-EST0048
85 287 6.1 561 10 B1156466 QVO-HT0036
86 286 6.1 431 14 CB852525 UI-CF-FN0
87 282 6.0 384 10 AW294086 UI-H-BL2-
88 281 6.0 529 9 AI051588 OZ02A09.X
89 281 6.0 656 10 BF033283 601454256
90 280 6.0 578 12 BM709052 UI-E-CQ1-
91 276 5.9 561 12 BM846750 K-EST0125
92 266 5.7 800 10 BF981057 602310275
93 263 5.6 309 10 AW405392 UI-HF-BL0
94 260 5.5 491 9 AI636360 TZ78C09.X
95 259 5.5 434 9 AI887805 TP97G07.X
96 258 5.5 477 14 H27205
97 258 5.5 887 13 BQ430213
98 257 5.5 338 10 BF915348 MR3-UT010
99 256 5.5 456 9 AI917072 TS74A01.X
100 250 5.3 441 9 AI362687 QK35C02.X

ALIGNMENTS

CF995807 759 bp mRNA linear EST 25-NOV-2003
AGENCOURT_16109164 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30708497 5', mRNA sequence.
CF995807
CF995807.1 GI:38512528
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 759)
NIH-MGC
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1074 row: m column: 18
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

source
1. .759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708497"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/notes="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. ligated with
EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. Average insert size
4-5kb. Adaptors 5' (AATTCGGCAGAGG)3' and 5'd
(CTCTGCGG)3'. 3' Linker sequence - GGGCGGCTGAGAGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTACCTCTCACTAAGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGG)3'. Library was constructed in the
Laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

ORIGIN

Query Match 13.3% Score 623; DB 14; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 563 CCGGAATTTCAGAGAAGCAGATGAAGTCTCTACAGAAAAGCAGAGCCAGCTGGTGAAGA 622
Db 8 CCGGAATTTCAGAGAAGCAGATGAAGTCTCTACAGAAAAGCAGAGCCAGCTGGTGAAGA 67
QY 623 GAAGGACCACTGGCGGTGAGCAGACAGCAGCGGCGTCTGCGCCGAGCAAGCTTCAGAG 682
Db 68 GAAGGACCACTGGCGGTGAGCAGACAGCAGCGGCGTCTGCGCCGAGCAAGCTTCAGAG 127
QY 683 CCTATGCGGTGAGCTGAGCGGCGCACAACCGCTCCCTCAAGGAAGAGGTGTGCGAGCGGC 742
Db 128 CCTATGCGGTGAGCTGAGCGGCGCACAACCGCTCCCTCAAGGAAGAGGTGTGCGAGCGGC 187
QY 743 CCGGAGGAGGAGAGAGAGCGGAAGAGGAGTGACCTGCACCTCCAGGTGACACTGAATGA 802
Db 188 CCGGAGGAGGAGGAGAGAGCGGAAGAGGAGTGACCTGCACCTCCAGGTGACACTGAATGA 247
QY 803 CATTGACCTGCAGATGAAGCAGACCAATGAGCGCAACTCCAACTGCGCCAGAGAACAT 862
Db 248 CATTGACCTGCAGATGAAGCAGACCAATGAGCGCAACTCCAACTGCGCCAGAGAACAT 307
QY 863 GGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGGAGGAGCATAT 922
Db 308 GGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGGAGGAGCATAT 367
QY 923 CGCAAAAGTCTTCAACACAGGACCTTACAAAGAGCTGTTGATGAGTGCCTCAAGCTCCAGCA 982
Db 368 CGCAAAAGTCTTCAACACAGGACCTTACAAAGAGCTGTTGATGAGTGCCTCAAGCTCCAGCA 427
QY 983 GGCCAGGAGATGCTAAAGAGGAGCAGAGAGCGGACCCAGCGGAGAAAGATTTCTCCT 1042
Db 428 GGCCAGGAGATGCTAAAGAGGAGCAGAGAGCGGACCCAGCGGAGAAAGATTTCTCCT 487
QY 1043 GAAAGAGGAGTATGAGTCCAGAGGATGTTGAGCTGATGAGTGAAGCAGCAGAGACCCACCT 1102
Db 488 GAAAGAGGAGTATGAGTCCAGAGGATGTTGAGCTGATGAGTGAAGCAGCAGAGACCCACCT 547
QY 1103 GAAGCAACAGCTTGCCTTATACACAGAGATTTGAGGAGTTCCAGAACACACTTTCCAA 1162
Db 548 GAAGCAACAGCTTGCCTTATACACAGAGATTTGAGGAGTTCCAGAACACACTTTCCAA 607
QY 1163 AAGCAGGAGTATTCACCAT 1185
Db 608 AAGCAGGAGTATTCACCAT 630

RESULT 2

BM800456 1085 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6419239 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534975
DEFINITION 5', mRNA sequence.
ACCESSION BM800456
VERSION BM800456.1 GI:19117279
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1085)
REFERENCE 1 (bases 1 to 1085)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12222 row: e column: 24
High quality sequence start: 30
High quality sequence stop: 679.
Location/Qualifiers
1. .1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:15534975"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN

Query Match 12.2%; Score 572; DB 12; Length 1085;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1185 TTCACGAGGAGATGAAAGATGACTAAGAGATCAAGAGCTGGAGAAACACACC 1244
DB 37 TTCAAGCAGGAGATGAAAGATGACTAAGAGATCAAGAGCTGGAGAAATACACACC 96
QY 1245 ATGTACCGGTCCCGTGGAGAGCAGCAACAGGCCCTGCTGAGATGGCTGAGGAGAA 1304
DB 97 ATGTACCGGTCCCGTGGAGAGCAGCAACAGGCCCTGCTGAGATGGCTGAGGAGAA 156
QY 1305 ACAGTCGGGATAAAGAACTGGAGCGCTGCGAGTAAATCCAAACGGCTGGAGAGCTG 1364
DB 157 ACAGTCGGGATAAAGAACTGGAGCGCTGCGAGTAAATCCAAACGGCTGGAGAGCTG 216
QY 1365 TCCTGGGCACTGCACAGAGCGCAATGACCTGACACAGAGGTACAGAGCTGAGTGCT 1424
DB 217 TCCTGGGCACTGCACAGAGCGCAATGACCTGACACAGAGGTACAGAGCTGAGTGCT 276
QY 1425 GGTGGCCAGGCTCCCTCACTCACAGTGGCCCTGAGAGAGGCCAGAGGGGCTGGGCT 1484
DB 277 GGTGGCCAGGCTCCCTCACTCACAGTGGCCCTGAGAGAGGCCAGAGGGGCTGGGCT 336
QY 1485 CAAGCACCAGCTCCCGAGGCTCACAGAGCGCTTGTACCCAGAGGACCCGAGCACA 1544
DB 337 CAAGCACCAGCTCCCGAGGCTCACAGAGCGCTTGTACCCAGAGGACCCGAGCACA 396
QY 1545 GAAGCATCAGGCAGACTGGGCTCAAGAGCCCACTCGCCAGGCGCTAGAGAGCTGG 1604
DB 397 GAAGCATCAGGCAGACTGGGCTCAAGAGCCCACTCGCCAGGCGCTAGAGAGCTGG 456
QY 1605 TGTGGGTATGCTGGGAGAGGAGCGGAGCCAGCCAGCGCTGCGCCCAATAAAGGCTCC 1664
DB 457 TGTGGGTATGCTGGGAGAGGAGCGGAGCCAGCCAGCGCTGCGCCCAATAAAGGCTCC 516
QY 1665 CATGCTGACGACCCATGCTCAAGCCAGGAGTTCCTGACCTGGCTGGCATCTGGCACT 1724
DB 517 CATGCTGACGACCCATGCTCAAGCCAGGAGTTCCTGACCTGGCTGGCATCTGGCACT 575
QY 1725 TGCATATTTGGATTTTGGTGGTCAAGTTTACGTACATAGGCAATTTTCAAGGCTTTCGA 1784
DB 576 TGCATATTTGGATTTTGGTGGTCAAGTTTACGTACATAGGCAATTTTCAAGGCTTTCGA 635
QY 1785 AATGCATTTATACCTGTAACTGTACAGTGGGCTTGCATTTGGGATGGGGTGTGTACAGA 1844

Db 636 AATGCATTTATACCTGTAACTGTACAGTGGGCTTGCATTTGNGGATGGGGTGTGTACAGA 695
QY 1845 TGAAGTCAGTGGCTTGTCTGTAGCTGAAGAGTCTTGAAGAGCGCTGTATCTGTAGCTG 1904
Db 696 TGAAGTCAGTGGCTTGTCTGTAGCTGAAGAGTCTTGAAGAGCGCTGTATCTGTAGCTG 755
QY 1905 CCATCAGTGAAGTGGCAGAGTGAAGTGAAGTCTTCTGTCTGATTTGAGGCTCAGA 1964
Db 756 CCATCAGTGAAGTGGCAGAGTGAAGTGAAGTCTTCTGTCTGATTTGAGGCTCAGA 815
QY 1965 CCCTCCCTCCCTCCCTT 1979
Db 816 CCCTCCCTCCCTCCCTT 830

RESULT 3
CD655044 835 bp mRNA linear EST 18-JUN-2003
LOCUS
DEFINITION
AGENCOURT_14551870 NIA Human H1 Embryonic Stem Cell cDNA library
(Long) Homo sapiens cDNA clone IMAGE:30428424 5', mRNA sequence.
CD655044
VERSION
CD655044.1 GI:31894941
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 835)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru KO
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM517 row: h column: 01
High quality sequence stop: 692.
Location/Qualifiers
1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30428424"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and HOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with TRIzol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558

FEATURES
source

(2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pGACTGCTGCTGATCGGAGCGGCGCCCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lox-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 11.8%; Score 556; DB 14; Length 835;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 CAGAACACACTTCCAAAGCAGCGAGGTATTACACATTCACAGCAGGATGGAAG 1205
DB |||||||
QY 1206 ATGACTAAGAAGATCAAGAGCTGGAGAAAGAACACCATGTACCGGTCCCGTGGAG 1265
DB |||||||
QY 1266 AGCAGAACAGCCCTGCTGATGCTGTAGGAAACAGTCCGGATAAAGACTG 1325
DB |||||||
QY 134 AGCAGAACAGCCCTGCTGATGCTGTAGGAAACAGTCCGGATAAAGACTG 193
QY 1326 GAGGGCTGCGAGTAAATCAACCGCTGGAGAGCTGTGCGGCGACTGCAGACAGAG 1385
DB |||||||
QY 1386 CGAATGACTGAACAGAGGTACAGACCTGAGTGTGCTGGCCAGGCTCCCTCACT 1445
DB |||||||
QY 1446 GACAGTGGCCCTGAGAGGAGCCAGAGGCGCTGGGCTCAAGACCCAGCTCCCGCAGG 1505
DB |||||||
QY 1506 GTCACAGAGCGCTTGTACCCAGAGCAGCCAGACAGAGCATCAGCCAGACTGG 1565
DB |||||||
QY 1566 CCTCAGAGCCCACTCCCGCAGGCTAGAGAGCTGTGGTCAATGCTGGGAGG 1625
DB |||||||
QY 1626 GAGCGCAGCCAGCGGCTGGCCATAAAGGCTCCCATGCTCAGCAGCCCAATGCT 1685
DB |||||||
QY 1686 GAAGCCAGATGTTCT 1701
DB |||||||

RESULT 4

LOCUS BU528395/c 895 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_1018232 NIH_MGC_101 Homo sapiens cDNA clone
IMAGE:6538337 5', mRNA sequence.
ACCESSION BU528395
VERSION BU528395.1 GI:22838836
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2702 row: 1 column: 17
High quality sequence stop: 564.

FEATURES

source
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6538337"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/notes="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 11.7%; Score 548; DB 13; Length 895;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 648; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4048 CCAAGATCTGTGTGTTCTTAGATTGCTAGTTCTCCAGGGAGCCACAGAGGTG 4107
DB |||||||
QY 4108 AGCTCAAGAGCGCATGGCTCTGCTAATAGTAATGTTTTCAGGGCTTGTCCAGCTGA 4167
DB |||||||
QY 4168 AGCTCAAGAGCGCATGGCTCTGCTAATAGTAATGTTTTCAGGGCTTGTCCAGCTGA 557
DB |||||||
QY 4168 GAGCTTCATGTCACCAAGATTCAGAGGTGTGAGAGGCTTTTTCAGGACTTTTTCAGTGA 4227
DB |||||||
QY 4228 TTGTTTTTCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTTCGTAGGAGACT 4287
DB |||||||
QY 496 TTGTTTTTCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTTCGTAGGAGACT 437
DB |||||||
QY 4288 GTTATTTCTGTAAGATGTTTAAACCTCTCCACCCCATCAGGTGGCCCTGAGGG 4347
DB |||||||
QY 436 GTTATTTCTGTAAGATGTTTAAACCTCTCCACCCCATCAGGTGGCCCTGAGGG 377
DB |||||||
QY 4348 CTGACCCGGAGGCCAGTGGAGCTGCTGTTGTCACGGGGAGGCCAAGCCCTGCTGAG 4407
DB |||||||
QY 376 CTGACCCGGAGGCCAGTGGAGCTGCTGTTGTCACGGGGAGGCCAAGCCCTGCTGAG 317
DB |||||||
QY 4408 CTGATTTCCAGTGTGCCCCCAGCCCTTCCGGCTTGCACAGCAGAGAGTGTCAACCC 4467
DB |||||||
QY 316 CTGATTTCCAGTGTGCCCCCAGCCCTTCCGGCTTGCACAGCAGAGAGTGTCAACCC 257
DB |||||||
QY 4468 AGGACAGCAGGAGCACTGCTCTCTTCCTTCCTGGGGGAAAGAGCTGCCCTTCGTC 4527
DB |||||||
QY 256 AGGACAGCAGGAGCACTGCTCTCTTCCTTCCTGGGGGAAAGAGCTGCCCTTCGTC 197
DB |||||||
QY 4528 CCTGTAAGCTCTTTCCTTATGGGCCCAACCCGGCCACTTCAGACTTGTGTAAGCTGCACTG 4587
DB |||||||

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 895)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2702 row: 1 column: 17

High quality sequence stop: 564.

source
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6538337"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/notes="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

/clone_lib="Human Retinal pigment epithelium/choroid cdna
(Un-normalized, unamplified): cs"
/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cdna
library in the pCMVSPORT6 vector was constructed at Life
Technologies (Rockville, MD; now part of Invitrogen Corp),
essentially following the protocols of the SuperScript
Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/>". The library code
designated was cs. For this library, cdna inserts were
cloned into the NotI/MluI sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."

ORIGIN		Query Match	11.1%;	Score 522;	DB 14;	Length 657;
		Best Local Similarity	99.7%;	Pred. No. 0;		
		Matches 622;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	4048	CCAGATCTGTGTTGTTCTTCTAGATTGCTAGCTTTCTCCAGGGACCCAGCAGGTG	4107			
DB	624	CCAAGATCTGTGTTGTTCTTCTAGATTGCTAGCTTTCTCCAGGGACCCAGCAGGTG	565			
QY	4108	AAGCTCAAGAGCCATGGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTGTCCAGCTGA	4167			
DB	564	AAGCTCAAGAGCCATGGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTGTCCAGCTGA	505			
QY	4168	GAGCTTCATGTCACCAAGATCTGAGAGGTGTCAGAGCAGCTTTTATTTATTTGTTG	4227			
DB	504	GAGCTTCATGTCACCAAGATCTGAGAGGTGTCAGAGCAGCTTTTATTTATTTGTTG	445			
QY	4228	TTTGTTTTCATGAGTTATCGGACCATGGCTGAGCTCAGGACCTTCTGTAGGAGACT	4287			
DB	444	TTTGTTTTCATGAGTTATCGGACCATGGCTGAGCTCAGGACCTTCTGTAGGAGACT	385			
QY	4288	GTATTTCTGTAAGATGTTATTTAAACCTCTCTCCACCCCATCGGTGGCCCTGAGGG	4347			
DB	384	GTATTTCTGTAAGATGTTATTTAAACCTCTCTCCACCCCATCGGTGGCCCTGAGGG	325			
QY	4348	CTACCCGGAGGCGCAGTCTGCTGTGTCACCGGGAGGGCCAGAGGCTGCTGAG	4407			
DB	324	CTACCCGGAGGCGCAGTCTGCTGTGTCACCGGGAGGGCCAGAGGCTGCTGAG	265			
QY	4408	CTGATTTCTCAGCTGTGCTCCACAGCCTTTCCGCTTSCACAGCAGAGTGTTCACCCC	4467			
DB	264	CTGATTTCTCAGCTGTGCTCCACAGCCTTTCCGCTTSCACAGCAGAGTGTTCACCCC	205			
QY	4468	AGGAGACCCAGCAGCTGCTCTCTTCTGCTTCTGGGGGAAAGAGCTGCTCTGTC	4527			
DB	204	AGGAGACCCAGCAGCTGCTCTCTTCTGCTTCTGGGGGAAAGAGCTGCTCTGTC	145			
QY	4528	CCTGTAACCTCTTCTTATGCGCCCAACCGGCACTCAGACTTGTGTAAGCTGCACTG	4587			
DB	144	CCTGTAACCTCTTCTTATGCGCCCAACCGGCACTCAGACTTGTGTAAGCTGCACTG	85			
QY	4588	GCAGCTTTTGTCTCTCTTGGGTATTCACACAGCCAGGAGCTGATTTGATGTAATTT	4647			
DB	84	GCAGCTTTTGTCTCTCTTGGGTATTCACACAGCCAGGAGCTGATTTGATGTAATTT	25			
QY	4648	TAAACCATTAATAAAGAGTCT 4671				
DB	24	TAAACCATTAATAAAGAGTCT 1				

RESULT 7
CA417053/c
LOCUS
DEFINITION
UI-H-PE0-bbp-n-03-0-UI.sl NCI CGAP_FEO Homo sapiens cdna clone
UI-H-PE0-bbp-n-03-0-UI 3', mRNA sequence.
ACCESSION
CA417053.1 GI:24779704

KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 591)
TITLE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Martin
cdna Library Preparation: Dr. M. Bento Soares, University of Iowa
cdna Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers
1. 591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-PE0-bbp-n-03-0-UI"
/tissue_type="Chondrosarcoma Cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_FEO"
/note="Vector: p77n3-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FEO is
a cDNA library containing the following tissue(s): a pool
of 3 chondrosarcoma cell lines (grade 2) The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p77n3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dr)18 tail. The sequence tag for this library is
CGTACCGAC. The cell lines was provided by Dr James Martin
of University of Iowa.
TAG_TISSUE=Human Grade 2 chondrosarcoma cell line pool
TAG_LIB=UI-H-PE0
TAG_SEQ=CGCTACGGAC"

ORIGIN		Query Match	10.4%;	Score 489;	DB 14;	Length 591;
		Best Local Similarity	99.7%;	Pred. No. 0;		
		Matches 589;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	4107	GAAGCTCAAGAGCCATGGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTGTCCAGCTG	4166			
DB	591	GAAGCTCAAGAGCCATGGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTGTCCAGCTG	532			
QY	4167	AGAGCTTCATGTCACCAAGATCTGAGAGGTGTCAGAGCAGCTTTTATTTATTTGTT	4226			
DB	531	AGAGCTTCATGTCACCAAGATCTGAGAGGTGTCAGAGCAGCTTTTATTTATTTGTT	472			
QY	4227	GTTTGTTCATGAGGTATTCGACCATGGCTGAGCTCAGGCATCTTCTGTAGGAGAC	4286			
DB	471	GTTTGTTCATGAGGTATTCGACCATGGCTGAGCTCAGGCATCTTCTGTAGGAGAC	412			
QY	4287	TGTTATTTCTGTAAGATGTTATTTAAACCTCTCCACCCCATCAGGTGGCCCTGAGG	4346			
DB	411	TGTTATTTCTGTAAGATGTTATTTAAACCTCTCCACCCCATCAGGTGGCCCTGAGG	352			
QY	4347	GCTACCCGGAGGCGCAGTGGAGCTGCTGTTGTCCACGGGGAGGGCCAAAGGCTGCTGA	4406			
DB	351	GCTACCCGGAGGCGCAGTGGAGCTGCTGTTGTCCACGGGGAGGGCCAAAGGCTGCTGA	292			

AGSCOURT:10437001 NIH_MGC_109 Homo sapiens CDNA clone
IMAGE:16598234 5', mRNA sequence.
BU849604
BU849604.1 GI:24034567
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutharia; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 799)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LICM2830 row: 1 column: 10
High quality sequence stop: 566.
Location/Qualifiers
1. .799
/organism="Homo sapiens"
FEATURES
source

4288 GTTATTCTGTAAGATGGTTATTTAAACCTCTCCACCCATCATCGGTGGCCCTGAGGG 4347
QY
Db 259 GTTATTCTGTAAGATGGTTATTTAAACCTCTCCACCCATCATCGGTGGCCCTGAGGG 318
QY 4348 CTGACCCGAGGCCAGTGGAGTGGCTGTGTCACGGGGGAGGCCAAGGCTCTCTGAG 4407
Db 319 CTGACCCGAGGCCAGTGGAGTGGCTGTGTCACGGGGGAGGCCAAGGCTCTCTGAG 378
QY 4408 CTGATTCTCCAGTGGTGGCCAGCCCTTCCCGCTTGCACACACAGAGGTGGTCAACCCC 4467
Db 379 CTGATTCTCCAGTGGTGGCCAGCCCTTCCCGCTTGCACACACAGAGGTGGTCAACCCC 438
QY 4468 AGGACAGCCAGGACCTCTCTCTTGGCCCTTCTGGGGGAA 4510
Db 439 AGGACAGCCAGGACCTCTCTCTTGGCCCTTCTGGGGGAA 481

RESULT 11

CB384172
LOCUS
DEFINITION
cDNA clone TgESTzh52a01.y1 5', mRNA sequence.

CB384172

ACCESSION

CB384172.1 GI:29133468

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajio, J.A., White, M.,
Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marr, M.,
Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,
Ritter, S., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I.,
Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Toxoplasma EST Project
Contact: Clifton, S.

Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
Fax: 314 286 1810

Email: toxo@wustl.edu

Contact David Sibley (toxost@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 562.

FEATURES

source

1. 597
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/db_xref="taxon:5811"
/clone="TgESTzh52a01.y1"
/dev_stage="3 day"
/lab_host="electroten Blue cells"
/clone_lib="TgME49 3 day invitro bradyzoite"
/note="Vector: pBluescript SK; Site 1: NotI; Site 2: SmaI;
cDNA was reverse transcribed using Superscript II and an
anchored (Pvobase) oligo dT primer, containing a NOTI
restriction site, according to standard protocols. Blunt
second strand cDNA was digested with NOTI restriction
endonuclease, purified and size fractionated with a
Sizedep 400 spun column (Amersham) and directionally
cloned into the SmaI and NotI sites of pBluescript SK.
Following electroporation into Electronen Blue cells
(Stratagene), 5 x 10⁶ primary CFU were bottle amplified in
semisolid culture media [2xLB+15% Seaprep agarose (BMA)].
After 48 hrs at 30 degrees Celcius, cells were harvested,
pooled and stored in 2xLB + 15% glycerol."

ORIGIN

Query Match

9.9%; Score 463; DB 14; Length 597;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4048 CCAGATCTCTGTGGTTCTTAGATTGCTAGCTTTCTCCAGGGGACACAGCAGGTG 4107
Db 115 CCAGATCTCTGTGGTTCTTAGATTGCTAGCTTTCTCCAGGGGACACAGCAGGTG 174
QY 4108 AAGCTCAAGAGCGCATGGCTCTGCTAATAGTAAATTTTTCAGGGGCTTTGTCAGCTGA 4167
Db 175 AAGCTCAAGAGCGCATGGCTCTGCTAATAGTAAATTTTTCAGGGGCTTTGTCAGCTGA 234
QY 4168 GAGCTTCATGTCACAGATCTGAGAGGTGTGACAGCAGCTTTTATTATTGTTG 4227
Db 235 GAGCTTCATGTCACAGATCTGAGAGGTGTGACAGCAGCTTTTATTATTGTTG 294
QY 4228 TTTGTTTTCATGAGGTATTCGGACCATGGCTGAGCTCAGGCACCTTCTGTAGGAGACT 4287
Db 295 TTTGTTTTCATGAGGTATTCGGACCATGGCTGAGCTCAGGCACCTTCTGTAGGAGACT 354
QY 4288 GTTATTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCATCATCGGTGGCCCTGAGGG 4347
Db 355 GTTATTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCATCATCGGTGGCCCTGAGGG 414
QY 4348 CTGACCCGAGGCCAGTGGAGTGGCTGTGTCACGGGGGAGGCCAAGGCTCTGCTGAG 4407
Db 415 CTGACCCGAGGCCAGTGGAGTGGCTGTGTCACGGGGGAGGCCAAGGCTCTGCTGAG 474
QY 4408 CTGATTCTCCAGTGTGTCGCCCGACCTTTCCGCTTGCACACACAGAGGTGGTCAACCCC 4467
Db 475 CTGATTCTCCAGTGTGTCGCCCGACCTTTCCGCTTGCACACACAGAGGTGGTCAACCCC 534
QY 4468 AGGACAGCCAGGACCTCTCTCTTGGCCCTTCTGGGGGAA 4510
Db 535 AGGACAGCCAGGACCTCTCTCTTGGCCCTTCTGGGGGAA 577

RESULT 12

LOCUS

DEFINITION

60771858F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858388 5',

ACCESSION

BG761234

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

817 bp mRNA linear EST 15-MAY-2001
mRNA sequence.
BG761234.1 GI:14071887
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 817)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI712 row: f column: 21
High quality sequence stop: 766.

FEATURES

source

1. 817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4858388"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:

ECORI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. "

ORIGIN

Query Match 9.8%; Score 462; DB 12; Length 817;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 612; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAAAAGAGCAGCCAGGACCAACCGAAGCAGGAGCCCGAGGAGCCCGCCAGCC 60
 Db 61 CAAAAGCAGCCAGGACCAACCGAAGCAGGAGCCCGAGGAGCCCGCCAGCC 120
 Qy 61 AGCGGGCTCTGAGTAGAAGCAGAGAGTCCCGGAGCAGCAGGCTCTCGAAGCCGG 120
 Db 121 AGCGGGCTCTGAGTAGAAGCAGAGAGTCCCGGAGCAGCAGGCTCTCGAAGCCGG 180
 Qy 121 AGCGGGCTCAAGCCAGACCGCTCAGTCTGGGGCCCTTCTGTATGTCCTGAGGAGCTGA 180
 Db 181 AGCGGGCTCAAGCCAGACCGCTCAGTCTGGGGCCCTTCTGTATGTCCTGAGGAGCTGA 240
 Qy 181 GCGCCCAACTGGAGACATATCTGAGCAGATATCTGTGTGACAAATACACAGGGGGCCCG 240
 Db 241 GCGCCCAACTGGAGACATATCTGAGCAGATATCTGTGTGACAAATACACAGGGGGCCCG 300
 Qy 241 GCGAGGATGGGCACAGGCTGACCGGCTGACCCGAGATGACAGAGAGTCCCGGACCT 300
 Db 301 GCGAGGATGGGCACAGGCTGACCGGCTGACCCGAGATGACAGAGAGTCCCGGACCT 360
 Qy 301 ATGTGGCAAGATGGGGAGCTTGACCACTCCAGTACTCTATGGAGAGAGAACCT 360
 Db 361 ATGTGGCAAGATGGGGAGCTTGACCACTCCAGTACTCTATGGAGAGAGAACCT 420
 Qy 361 CCAAGGGGATCCAAACAGAGAGATCGGAGAGTACAGAGTCCGAGAGCCAGGAC 420
 Db 421 CCAAGGGGATCCAAACAGAGAGATCGGAGAGTACAGAGTCCGAGAGCCAGGAC 480
 Qy 421 ATCGAAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Db 481 ATCGAAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Qy 481 TGCAAGATGAAATATCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Db 541 TGCAAGATGAAATATCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Qy 541 ATGCTGAATCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 601 ATGCTGAATCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 Qy 601 AGCAGAGCCAGCTGG 615
 Db 661 AGCAGAGCCAGCTGG 675

RESULT 13

BU732511/c
 LOCUS 713 bp mRNA linear EST 09-OCT-2002
 DEFINITION UI-E-CQ1-afx-h-04-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
 UI-E-CQ1-afx-h-04-0-UI 3', mRNA sequence.
 ACCESSION BU732511
 VERSION BU732511.1 GI:23658483
 KEYWORDS Est.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 713)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 FORWARD

POLYA=res.

Location/Qualifiers

1..713

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/c1ones="UI-E-CQ1-afx-h-04-0-UI"

/tissue_type="optic nerve"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-CQ1"

/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CQ1 is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCATTAGTG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG_TISSUE=human optic nerve

TAG_LIB=UI-E-CQ1

TAG_SEQ=CCATTAGTG

ORIGIN

Query Match 9.6%; Score 451; DB 13; Length 713;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 551; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4145 TTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACCAGATTCGAGAGGTGTGACGA 4204
 Db 553 TTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACCAGATTCGAGAGGTGTGACGA 494
 Qy 4205 GCATTTTTTTTTTTTATTTCTGTTTCTTCCATGAGTTATCGGACCATGGGCTGAGC 4264
 Db 493 GCATTTTTTTTTTTTATTTCTGTTTCTTCCATGAGTTATCGGACCATGGGCTGAGC 434
 Qy 4265 TCAGGCATTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTATTAACCTCTCTCCA 4324
 Db 433 TCAGGCATTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTATTAACCTCTCTCCA 374
 Qy 4325 CCCCATCAGGTGGCCCTGAGGCTGACCCGGAGGCGCAGTGGAGCTGCTGGTGCACG 4384
 Db 373 CCCCATCAGGTGGCCCTGAGGCTGACCCGGAGGCGCAGTGGAGCTGCTGGTGCACG 314
 Qy 4385 GGGGAGGCGCAAGGCTGCTGAGCTGATTTCTCCAGCTGTGSCCCAGCCTTTCCGCTTG 4444

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8895548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University
cDNA Library Arrayed by: Dr. M. Bento Soares, University
cDNA Sequencing by: Dr. M. Bento Soares, University of
Clone Distribution: Researchers may obtain clones from
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
polyA=yes.

FEATURES

```

Location/Qualifiers
1. 598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CQ1-aex-h-16-0-UI"
/tissue_type="optic nerve"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CQ1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoG I; Site 2: Not I; UI-E-CQ1 is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-805, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCATTAACTG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
TAG TISSUE=human
TAG LIB=UI-E-CQ1
TAG_SEQ=CCATTAACTG

```

ORIGIN

Query Match 9.5%; Score 445; DB 13; Length 598;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 595; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

4100	AGCAGGTCGAGCTCAAGAGCGCATGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTTGT	QY
598	AGCAGGTCGAGCTCAAGAGCGCATGCTCTGCTAATAGTAAATGTTTTCAGGGCCCTTGT	Db
4160	CCAGCTGAGAGCTTCATGTCCTCACACAGATCTCAGAGGTTCTCAGCAGCACTTTTTTTTTTT	QY
538	CCAGCTGAGAGCTTCATGTCCTCACACAGATCTCAGAGGTTCTCAGCAGCACTTTTTTTTTTT	Db
4220	ATTTGTTGTTGTTTTCATGAGGTTATCGGACCATGGGCTGAGCTCAGGCACACTTCTGT	QY
478	ATTTGTTGTTGTTTTCATGCGGTTATCGGACCATGGGCTGAGCTCAGGCACACTTCTGT	Db
4280	AGGAGACTGTTTATTTCTGTAAGAGATGGTTATTAAACCTCTCTCCACCCCATCACGGTGGC	QY
418	AGGAGACTGTTTATTTCTGTAAGAGATGGTTATTAAACCTCTCTCCACCCCATCACGGTGGC	Db
4340	CCTGAGGGCTGACCCGAGGCGCCAGTGGAGCTGCTGGTGTCCAGGGGGGAGGGCCAAAGC	QY
358	CCTGAGGGCTGACCCGAGGCGCCAGTGGAGCTGCTGGTGTCCAGGGGGGAGGGCCAAAGC	Db
299	CCTGAGGGCTGACCCGAGGCGCCAGTGGAGCTGCTGGTGTCCAGGGGGGAGGGCCAAAGC	QY

Qy	4400	CTGCTGAGCTGATTTCTTCAGACTGCTGCCCCAGGCTTTCCGCTTGCACAGCAGCAGAGGTG	4455
Db	298	GTGCTGAGCTGATTTCTTCAGACTGCTGCCCCAGGCTTTCCGCTTGCACAGCAGAGGTG	239
Qy	4460	GTACACCCAGGACAGCAGGACGCTGCTCTCTTCGCTTCTCTGGGGGAAAGGACTGC	4519
Db	238	GTACACCCAGGACAGCAGGACGCTCTCTCTTCGCTTCTCTGGGGGAAAGGACTGC	179
Qy	4520	CTTCTGTCCCTGTAACCTGCTTTCTTATGGCCCAACCCGGGCACCTCAGACTTGTTTGAAG	4579
Db	178	CTTCTGTCCCTGTAACCTGCTTTCTTATGGCCCAACCCGGGCACCTCAGACTTGTTTGAAG	119
Qy	4580	CTGCACTGCAGCTTTTTTGTCTCCTTTGGGTATTCACACAGCCAGGACTTGATTTTG	4639
Db	118	CTGCACTGCAGCTTTTTTGTCTCCTTTGGGTATTCACACAGCCAGGACTTGATTTTG	59
Qy	4640	ATGTATTTTTAAACCACTTAAATAAAGACTCTGTTCCTTTAAAAAATAAAAAAAAAAAAA	4697
Db	58	ATGTATTTTTAAACCACTTAAATAAAGACTCTGTTCCTTTAAAAAATAAAAAAAAAAAAA	1

RESULT 16
 BQ896779/c
 LOCUS
 DEFINITION
 BQ896779
 ACCESSION
 BQ896779
 VERSION
 BQ896779.1
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 920 bp mRNA linear EST 16-AUG-2002
 AGSCENTR08715768 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425830
 5', mRNA sequence.
 NIH-MGC <http://mgc.nci.nih.gov/>.
 1 (bases 1 to 920)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgavsb-r@mail.nih.gov

FEATURES
SOURCE

```

1. 520
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IWAGS:6425830"
/tissue_type="epidermoid carcinoma, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site: 1: EcoRI
XhoI; cDNA made by oligo-dT priming. Direction:
into EcoRI/XhoI sites using the following 5' and
GGCACGAG(G). Library constructed by Ling Hong
laboratory of Gerald M. Rubin (University of
Berkeley) using ZAP-cDNA synthesis kit (Stratagene)
Superscript II RT (Life Technologies). Note: this
NIH MGC library."

```

ORIGIN

```
Query Match          9.5%; Score 444; DB 13; Length 920;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 594; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

OV 4059 GTTGGTTTCTTAGATTGCTAGCTTTTCCCTCCAGGGACCACAGCAGGTGAAGCTCAAGAG 4118

Db 617 GTTGGTTCTTAGATGCTAGCTTTCTCCAGGGGACCAACAGAGGTGAAGCTCAAG 558
QY 4119 CCGATGGCTCTCTAATAGTAATTTGTTTTCAGGCGCTTGTCCAGCTGAGAGCTTCATGT 4178
Db 557 CCGATGGCTCTCTAATAGTAATTTGTTTTCAGGCGCTTGTCCAGCTGAGAGCTTCATGT 498
QY 4179 CCACAGATTCGAGAGGTGTCAGAGACACTTTTTTTTTTTTATTTGTTGTTTTCCTCA 4238
Db 497 CCACAGATTCGAGAGGTGTCAGAGACACTTTTTTTTTTTTATTTGTTGTTTTCCTCA 438
QY 4239 TGAGGTATTCGAGACCTGGGCTGAGCTCAGGACACTTTCTGTAGGAGACTGTTATTTCTGT 4298
Db 437 TGAGGTATTCGAGACCTGGGCTGAGCTCAGGACACTTTCTGTAGGAGACTGTTATTTCTGT 378
QY 4299 AAAGATGTTATTAACCTCTCCACCCATCAGGTGGCCCTGAGGCTGACCCGAG 4358
Db 377 AAAGATGTTATTAACCTCTCCACCCATCAGGTGGCCCTGAGGCTGACCCGAG 318
QY 4359 GCCAGTGGAGCTGCCCTGTGTCTCAGGGGGAGGCGCAAGGCTGTGAGCTGATTCTCCA 4418
Db 317 GCCAGTGGAGCTGCCCTGTGTCTCAGGGGGAGGCGCAAGGCTGTGAGCTGATTCTCCA 258
QY 4419 GCTGTCGCCCGAGCTTTCCGCTTGCACAGACAGAGTGTGTACCCAGGAGCAGCCA 4478
Db 257 GCTGTCGCCCGAGCTTTCCGCTTGCACAGACAGAGTGTGTACCCAGGAGCAGCCA 198
QY 4479 GGCACCTGCTCTCTTGGCCCTTCTCGGGGAAAGAGCTGCTTCTGCTCCCTGTAACCTGC 4538
Db 197 GGCACCTGCTCTCTTGGCCCTTCTCGGGGAAAGAGCTGCTTCTGCTCCCTGTAACCTGC 138
QY 4539 TTTCTCTATGGCCCAACCCGGGCACTCAGACTGTTTGAAGCTGCACTGGCAGCTTTT 4598
Db 137 TTTCTCTATGGCCCAACCCGGGCACTCAGACTGTTTGAAGCTGCACTGGCAGCTTTT 78
QY 4599 GTCTCTGTTGGTATTCAACAGCAGGAGCTGCTGATTTGTATTTTAAACAC 4655
Db 77 GTCTCTGTTGGTATTCAACAGCAGGAGCTGCTGATTTGTATTTTAAACAC 21

RESULT 17
BF939517/c 545 bp mRNA linear EST 22-JAN-2001
LOCUS rac77f09.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3440416
DEFINITION 3', mRNA sequence.
ACCESSION BF939517
VERSION BF939517.1 GI:12356837
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
REFERENCE NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute / National Institute of Neurological
TITLE Disorders and Stroke Brain Tumor Genome Anatomy Project
(CGAP/BSGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 471.
Location/Qualifiers
1. 545

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3440416"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 9.4%; Score 443; DB 10; Length 545;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 543; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4135 TAGTAAATTTGTTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACAGATTTCTGAGA 4194
Db 545 TAGTAAATTTGTTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACAGATTTCTGAGA 486
QY 4195 GGTGTACGACGACACTTTTTTTTATTTGTTGTTTCTTTTCCATGAGGTATTCGGACCA 4254
Db 485 GGTGTACGACGACACTTTTTTTTATTTGTTGTTTCTTTTCCATGAGGTATTCGGACCA 426
QY 4255 TGGGTGAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAAGATGGTTATTAA 4314
Db 425 TGGGTGAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAAGATGGTTATTAA 366
QY 4315 CCCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCT 4374
Db 365 CCCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCT 306
QY 4375 GGTGTCCACGGGGAGGCGCCAGGCCCTGCTGAGCTGATTTCTCCAGCTGCTGCCCGAGCT 4434
Db 305 GGTGTCCACGGGGAGGCGCCAGGCCCTGCTGAGCTGATTTCTCCAGCTGCTGCCCGAGCT 246
QY 4435 TTCCGCTTTCACAGCAGCAGAGGTGGTTCACCCAGGACAGCCAGGACACCTGCTCTCTT 4494
Db 245 TTCCGCTTTCACAGCAGCAGAGGTGGTTCACCCAGGACAGCCAGGACACCTGCTCTCTT 186
QY 4495 GCCTTCTCTGGGGAAAGAGGAGCTCCCTTCTGTCCTGTAACCTGCTTTTATGGCCCAA 4554
Db 185 GCCTTCTCTGGGGAAAGAGGAGCTCCCTTCTGTCCTGTAACCTGCTTTTATGGCCCAA 126
QY 4555 CCCTGCTTTCAGACTGTTTGAAGCTGCACTGGCAGCTTTTGTCTCTTTCGGTATT 4614
Db 125 CCCTGCTTTCAGACTGTTTGAAGCTGCACTGGCAGCTTTTGTCTCTTTCGGTATT 66
QY 4615 CACACAGCCAGGAGCTGATTTTGTATGATTTTAAACCATTTAAAGAGTCTGTT 4674
Db 65 CACACAGCCAGGAGCTTGTATTTGTATGATTTTAAACCATTTAAAGAGTCTGTT 6
QY 4675 GCCTT 4679
Db 5 GCCTT 1

RESULT 18
BU506849

LOCUS BU506849 880 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10302113 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6499924
5', mRNA sequence.
ACCESSION BU506849
VERSION BU506849.1 GI:22813082
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 880)
 NHT-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M14051 row: p column: 05
 High quality sequence stop: 640.
 Location/Qualifiers
 1..880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6499924"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-Sport6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."
 ORIGIN
 Query Match 9.4%; Score 443; DB 13; Length 880;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 3575 TTTGCTAGGCGAGAGGTCAGGAGAGGAGGAGGATGACCTTCCCTTGAAGAG 3634
 Db 17 TTTGCTAGGCGAGAGGTCAGGAGAGGAGGAGGATGACCTTCCCTTGAAGAG 76
 3635 AGGGAAAGGCTACAGTGGCCACATATTCCTGACTCACATTCAGTACCTCTTAA 3694
 Db 77 AGGGAAAGGCTACAGTGGCCACATATTCCTGACTCACATTCAGTACCTCTTAA 136
 3695 TGCTGTGAGGAGCTGAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 3754
 Db 137 TGCTGTGAGGAGCTGAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 196
 3755 GCAGGTGGCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCAAGGAGGAGCC 3814
 Db 197 GCAGGTGGCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCAAGGAGGAGCC 256
 3815 TTTGCTCAGAGTCCATGCTAGAGTCTAGAGTCTGAGGAGGAGGAGGAGGAGG 3874
 Db 257 TTTGCTCAGAGTCCATGCTAGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGG 316
 3875 CTTTTTCTCTCTTCAATTGACCAATCAACCAATCACTACAGTGTCTGCTTCTGCT 3934
 Db 317 CTTTTTCTCTCTTCAATTGACCAATCAACCAATCACTACAGTGTCTGCTTCTGCT 376
 3935 TCCAAAGTACCCAGGCTGGGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3994
 Db 377 TCCAAAGTACCCAGGCTGGGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGG 436
 3995 CAGTGTCTTCTCTCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4017
 Db 437 CAGTGTCTTCTCTCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 459
 RESULT 19
 BE857092/c
 LOCUS
 DEFINITION 521 bp mRNA linear EST 29-SEP-2000
 7g24b10.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3307387 3',
 mRNA sequence.

ACCESSION BE857092.1 GI:10370773
 VERSION BE857092.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 521)
 NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurological
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/STGAP), Tumor Gene Index
 Unpublished (1998)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 473.
 Location/Qualifiers
 1..521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3307387"
 /tissue_type="glioblastoma (pooled)"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Brn23"
 /note="Organ: Brain; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."
 ORIGIN
 Query Match 8.9%; Score 419; DB 10; Length 521;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 519; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 4159 TCCAGCTGAGAGCTTCATGTCACCCAGATTCGAGAGGTCTCAGCAGCACTTTT 4218
 Db 521 TCCAGCTGAGAGCTTCATGTCACCCAGATTCGAGAGGTCTCAGCAGCACTTTT 462
 4219 TATTGTGTTGTTTTCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTCTG 4278
 Db 461 TATTGTGTTGTTTTCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTCTG 402
 4279 TAGGAGACTGTATTCTGTAAGATGTTATTAACTCTCCACCCCATCAGGTTG 4338
 Db 401 TAGGAGACTGTATTCTGTAAGATGTTATTAACTCTCTCCACCCCATCAGGTTG 342
 4339 CCTGAGGCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4398
 Db 341 CCTGAGGCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 282
 4399 CTTGCTGAGTGTATCTCCAGCTGTGCCCCAGCTTTCGCTTTCACAGCAGAGGT 4458
 Db 281 CTTGCTGAGTGTATCTCCAGCTGTGCCCCAGCTTTCGCTTTCACAGCAGAGGT 222
 4459 GGTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4518
 Db 221 GGTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162

```
QY 4519 CTTCTGTCCTCTACTCTCTTCTCTTATGAGCCCAAGCCGCTCAGCTGTTGTTGAA 4578
Db 161 CTTCTGTCCTCTACTCTCTTCTCTTATGAGCCCAAGCCGCTCAGCTGTTGTTGAA 102
QY 4579 GTGCACTGCGAGCTTTTGTCTCTTTGGTATTACAAAGCCAGGAGCTTGAATTT 4638
Db 101 GTGCACTGCGAGCTTTTGTCTCTTTGGTATTACAAAGCCAGGAGCTTGAATTT 42
QY 4639 GATGATTTTAAACCAATTAATAAAGAGCTGTTGCTT 4679
Db 41 GATGATTTTAAACCAATTAATAAAGAGCTGTTGCTT 1

RESULT 20
LOCUS BG827157 932 bp mRNA linear EST 22-MAY-2001
DEFINITION 60251143F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903690 5',
mRNA sequence.
ACCESSION BG827157
VERSION BG827157.1 GI:14174744
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1802 row: f column: 11
High quality sequence start: 6
High quality sequence stop: 775.
FEATURES
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4903690"
/tissue_type="thadomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_17"
/notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
QY 842 CAAGCTGGCCCAAGAGAACTGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTA 901
Db 125 CAAGCTGGCCCAAGAGAACTGGAGCTGGCTGAGAGGCTCAGAGAGCTGATTGAGCAGTA 184
QY 902 TGAGCTGGCGAGGAGCATATCGACAAGTCTTCAAAACAAGGACCTACAACAGCAGCT 961
Db 185 TGAGCTGGCGAGGAGCATATCGACAAGTCTTCAAAACAAGGACCTACAACAGCAGCT 244
QY 962 GGTGGATGCCAAGC-TCCAGCAGSCCCAGGAGATGCTAAAGGAGGCGAGAGAGCGGCACC 1020
Db 245 GGTGGATGCCAAGCTTCCAGCAGSCCCAGGAGATGCTAAAGGAGGCGAGAGAGCGGCACC 304
QY 1021 AGCGGAGAAAGGATTTTCTCTGAAAGAGGCGAGTAGTCCCAGAGATGTGTGAGCTGA 1080
Db 305 AGCGGAGAAAGGATTTTCTCTGAAAGAGGCGAGTAGTCCCAGAGATGTGTGAGCTGA 364
QY 1081 TGAAGCAGCAAGAGACCCACCTGAGCAACAGCTTGCCTTATACACAGAGAGTTTTCAGG 1140
Db 365 TGAAGCAGCAAGAGACCCACCTGAGCAACAGCTTGCCTTATACACAGAGAGTTTTCAGG 424
QY 1141 AGTTCCAGAACACACTTTTCCAAAAGCAGCAGGATTTTCCACACATTTCAAGCAGGAGATGG 1200
Db 425 AGTTCCAGAACACACTTTTCCAAAAGCAGCAGGATTTTCCACACATTTCAAGCAGGAGATGG 484
QY 1201 AAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACACCATATCGGTCCCGT 1260
Db 485 AAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACACCATATCGGTCCCGT 544

RESULT 21
LOCUS BG928861 941 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT 10013724 NIH_MGC_101 Homo sapiens cDNA clone
IMAGE:6497616 5', mRNA sequence.
ACCESSION BG928861
VERSION BG928861.1 GI:22343892
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 941)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2678 row: p column: 01
High quality sequence stop: 558.
FEATURES
Location/Qualifiers
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6497616"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

```

ORIGIN
Query Match      8.9%; Score 419; DB 13; Length 941;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3575 TTTGCTAGGCGAGAGAGGTGTCAGGGAAGGAGGAGGTATGACCTTCCCTTGAAGAG 3634
      |||||
Db 179 TTTGCTAGGCGAGAGAGGTGTCAGGGAAGGAGGAGGTATGACCTTCCCTTGAAGAG 238
      |||||
QY 3635 AGGGGAAGGCTTACAGTGGCCACATAATTCCTGACTCACATTCAGTACCTCTTAA 3694
      |||||
Db 239 AGGGGAAGGCTTACAGTGGCCACATAATTCCTGACTCACATTCAGTACCTCTTAA 298
      |||||
QY 3695 TGCTGTGAGGAGCTGAGAGCTCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 3754
      |||||
Db 299 TGCTGTGAGGAGCTGAGAGCTCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 358
      |||||
QY 3755 GCAGGTGGCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCACGGAAGGTGAGCC 3814
      |||||
Db 359 GCAGGTGGCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCACGGAAGGTGAGCC 418
      |||||
QY 3815 TTTGCTCAGAGCTCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTA 3874
      |||||
Db 419 TTTGCTCAGAGCTCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCCTA 478
      |||||
QY 3875 CTTTTTCTCTCTCATGACCAATCAACCAATCACTACAGTGTCTGTCTGTGTT 3934
      |||||
Db 479 CTTTTTCTCTCTCATGACCAATCAACCAATCACTACAGTGTCTGTCTGTGTT 538
      |||||
QY 3935 TCCAAAGTACCCAGCTCTGGCCAGATGCGGGAGGTGCTATCCATGATGAGG 3993
      |||||
Db 539 TCCAAAGTACCCAGCTCTGGCCAGATGCGGGAGGTGCTATCCATGATGAGG 597
      |||||

RESULT 22
BI260369 724 bp mRNA linear EST 17-JUL-2001
LOCUS 60295420F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108788 5',
DEFINITION mRNA sequence.
ACCESSION BI260369
VERSION BI260369.1 GI:14818605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 724)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11263 row: h column: 05
High quality sequence stop: 719.
Location/Qualifiers
1..724
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5108788"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT."

ORIGIN
Query Match      8.9%; Score 417; DB 12; Length 724;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4092 GGGACCCAGCAGAGGTGAAGCTCAAGAGCGCATGCTCTGCTAATAGTAAATTTGTTT 4151
      |||||
Db 184 GGGACCCAGCAGAGGTGAAGCTCAAGAGCGCATGCTCTGCTAATAGTAAATTTGTTT 243
      |||||
QY 4152 GGCCTTGTGAGTGAAGCTTCAATGTCACAGATCTCAGAGGTGTCAGAGCACTTT 4211
      |||||
Db 244 GGCCTTGTGAGTGAAGCTTCAATGTCACAGATCTCAGAGGTGTCAGAGCACTTT 303
      |||||
QY 4212 TTTTATTTTATTTGTTTGTTCATGAGGTATTCGGACCATGGGCTGAGCTCAGGCA 4271
      |||||
Db 304 TTTTATTTTATTTGTTTGTTCATGAGGTATTCGGACCATGGGCTGAGCTCAGGCA 363
      |||||
QY 4272 CTTTCTGTAGAGACTTTATTTCTGTAAAGATGGTTATTTAAACCTCTCTCCACCCATC 4331
      |||||
Db 364 CTTTCTGTAGAGACTTTATTTCTGTAAAGATGGTTATTTAAACCTCTCTCCACCCATC 423
      |||||
QY 4332 ACGTGTGCTGAGGCTGACCGGAGGCCAGTGGAGTGCCTGGTGTCCACGGGGAGG 4391
      |||||
Db 424 ACGTGTGCTGAGGCTGACCGGAGGCCAGTGGAGTGCCTGGTGTCCACGGGGAGG 483
      |||||
QY 4392 GCCAAGGCTGCTGAGTGTATTTCCAGTGTCTGCCACCCCTTCCGCTTCCACAGCA 4451
      |||||
Db 484 GCCAAGGCTGCTGAGTGTATTTCCAGTGTCTGCCACCCCTTCCGCTTCCACAGCA 543
      |||||
QY 4452 CAGAGTGTGTCACCCAGGAGCAGCAGCAGTGTCTCTTCCCTTCCCTGGGGG 4508
      |||||
Db 544 CAGAGTGTGTCACCCAGGAGCAGCAGCAGTGTCTCTTCCCTTCCCTGGGGG 600
      |||||

RESULT 23
BI194071 890 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT_7932633 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6157145
DEFINITION 5', mRNA sequence.
ACCESSION BI194071
VERSION BI194071.1 GI:22708055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 890)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13502 row: i column: 18
High quality sequence stop: 643.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6157145"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"

FEATURES
source

```

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 8.7%; Score 410; DB 13; Length 890;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 95 CAGCAGCCAGGCTCTCGAAGCCGAGGGGCTCAAGCAGAAAGCGTCTAGTCTGGGGC 154
Db 315 CAGCAGCCAGGCTCTCGAAGCCGAGGGGCTCAAGCAGAAAGCGTCTAGTCTGGGGC 374
QY 155 CTTCTGTGATCTCTGAGGAGCTGAGCGCCCAACTGGAAGACATACCTGAGACATACCTG 214
Db 375 CTTCTGTGATCTCTGAGGAGCTGAGCGCCCAACTGGAAGACATACCTGAGACATACCTG 434
QY 215 TGTGACATATACCGAGGGGGCCCGGAGGATGGGCACAGGCTGAGCGGCTGAACC 274
Db 435 TGTGACATATACCGAGGGGGCCCGGAGGATGGGCACAGGCTGAGCGGCTGAACC 494
QY 275 CGAAGATGAGAGAGTCTCCGAGACCTATGTGCAAGGAATGGGAGCCCTGAACCACTCC 334
Db 495 CGAAGATGAGAGAGTCTCCGAGACCTATGTGCAAGGAATGGGAGCCCTGAACCACTCC 554
QY 335 AGTAGTCTATGAGAGAGAGCAACCTCCAAAGGGGATCCAAACACAGAGAGATCCGGCA 394
Db 555 AGTAGTCTATGAGAGAGAGCAACCTCCAAAGGGGATCCAAACACAGAGAGATCCGGCA 614
QY 395 GAGTGAAGGCTCGAGACCGAGACCATCGAAGGCCACAGGAGAGAAAGCAAGGG 454
Db 615 GAGTGAAGGCTCGAGACCGAGACCATCGAAGGCCACAGGAGAGAAAGCAAGGG 674
QY 455 TTTGGGAGGAGATCAGCTTCTGATGAGACATGATGATCTGAGTACCTGAGTACCTGAGG 514
Db 675 TTTGGGAGGAGATCAGCTTCTGATGAGACATGATGATCTGAGTACCTGAGTACCTGAGG 734
QY 515 GAAGTGGCTGCTGTGTGCAAGAGATGCTGAACTGCTGG 555
Db 735 GAAGTGGCTGCTGTGTGCAAGAGATGCTGAACTGCTGG 775

RESULT 24

BQ643746
LOCUS 1061 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT 8341610 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6267874
5', mRNA sequence.
ACCESSION BQ643746
VERSION BQ643746.1 GI:21767918
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCW438 row: 0 column: 11
High quality sequence start: 2
High quality sequence stop: 519.
Location/Qualifiers
1..1061

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6267874"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_100"
/notes="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming, directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.9kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 8.7%; Score 408; DB 13; Length 1061;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 528; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2524 GCCATCTTGATGCTGAGGTTTCCTGTTGGTGAGATCAGGTTGTTGGTAAAGAAAG 2583
Db 22 GCCATCTTGATGCTGAGGTTTCCTGTTGGTGAGATCAGGTTGTTGGTAAAGAAAG 81
QY 2584 GAAAGGCTCTGATGCTGTTGGCACAAGCTTACCTGTTGGGTTTCAGTCTCGAGAGCCCA 2643
Db 82 GAAAGGCTCTGATGCTGTTGGCACAAGCTTACCTGTTGGGTTTCAGTCTCGAGAGCCCA 141
QY 2644 CCACAGTTCCTCATCAGACATGCTCCATGACAGAGTGTGGTCCCATGTCAGTGC 2703
Db 142 CCACAGTTCCTCATCAGACATGCTCCATGACAGAGTGTGGTCCCATGTCAGTGC 201
QY 2704 CTCTTTGGCTTCATGGTCTTTCTGCTTCCTGCCCCACCCACACATGTGCAATCTCTCAA 2763
Db 202 CTCTTTGGCTTCATGGTCTTTCTGCTTCCTGCCCCACCCACACATGTGCAATCTCTCAA 261
QY 2764 GATTGTCTGATTCATTTCTGGGACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2823
Db 262 GATTGTCTGATTCATTTCTGGGACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
QY 2824 CCTGTGGG-GCCCATAGCTGTTCTTAACAGTAAGAAATGAATGAATGAATGAATGAAT 2882
Db 322 CCTGTGGGAGCCCATAGCTGTTCTTAACAGTAAGAAATGAATGAATGAATGAATGAAT 381
QY 2883 GGGCCCCAGAAATCCATAAAATGCTGCAGACAGTGTGTTCTGTGCTCTGTTTACCCCC 2942
Db 382 GGGCCCCAGAAATCCATAAAATGCTGCAGACAGTGTGTTCTGTGCTCTGTTTACCCCC 441
QY 2943 ACTCCAGTACATACTACTGATGCTGTAGAGCCATTCATATGCTGATGCTGATGCTGCT 3002
Db 442 ACTCCAGTACATACTACTGATGCTGTAGAGCCATTCATATGCTGATGCTGATGCTGCT 501
QY 3003 GTTGCAAACTGCCAGGGTATTAGCCAGTGTGTTGTCAGAGCAGTTTC 3051
Db 502 GTTGCAAACTGCCAGGGTATTAGCCAGTGTGTTGTCAGAGCAGTTTC 550

RESULT 25

BQ692313
LOCUS 773 bp mRNA linear EST 22-DEC-2000
DEFINITION 602249206F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334352 5',
mRNA sequence.
ACCESSION BQ692313
VERSION BQ692313.1 GI:11977721
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 773)
NIH-MGC <http://mgi.nci.nih.gov/>.

ACCESSION	CA313019	CA313019.1	GI:24531117	
VERSION	EST.			
KEYWORDS				
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 543)			
TITLE	Bonaldo, M.F., Lennon, G. and Soares, M.B.			
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery			
MEDLINE	Genome Res. 6 (9), 791-806 (1996)			
PUBMED	97044477			
COMMENT	8889548			
	Contact: McCray, PB			
	University of Iowa			
	2024 University of Iowa Med Labs, Iowa City, IA 52242, USA			
	Tel: 319 356 4866			
	Fax: 319 356 7171			
	Email: paul-mccray@uiowa.edu			
	Tissue Procurement: Dr. M. J. Welsh, University of Iowa			
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa			
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa			
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa			
	Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com) or from Open Biosystems (www.openbiosystems.com).			
	The following repetitive elements were found in this cDNA sequence: 1-55, >POLY_A#Simple_repeat (matched complement) 500-521, >POLY_A#Simple_repeat			
	Seq primer: M13 FORWARD			
	POLYA=Yes.			
FEATURES	Location/Qualifiers			
source	1..543			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="UI-CF-FNO-aex-c-16-0-UI"			
	/tissue="Human Lung Epithelial cells"			
	/lab_host="DH10B (Life Technologies) (TI phage resistant)"			
	/clone_libs="UI-CF-FNO"			
	/note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalised Human lung epithelial cell libraries (EN1 and DUL) The library was subtracted according to according to Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bentso-soares@uiowa.edu			
	TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368			
	TAG LIB=UI-CF-FNO			
	TAG_SEQ=GGCTGTAGGC			
ORIGIN				
Query Match	8.2%; Score 387; DB 14; Length 543;			
Best Local Similarity	99.6%; Pred. No. 0;			
Matches 487; Conservative	0; Mismatches 2; Indels 0; Gaps 0;			
QY	4209 TTTTATTTTATTTGTTGTTTTCATGAGTTATCGGACCATGGCTGAGCTCAG	4268		
DB	29 TTTTATTTTATTTGTTGTTTTCATGAGTTATCGGACCATGGCTGAGCTCAG	88		
QY	4269 GCATTTCTGTAGGAGACTGTTATTTCTGAAGATGGTTATTAAACCTCTCCACCC	4328		
DB	89 GCATTTCTGTAGGAGACTGTTATTTCTGAAGATGGTTATTAAACCTCTCCACCC	148		
QY	4329 ATCAGGTGGCCCTGAGGCTGACCCGAGCCAGTGCCTGGTGTCCACGGGG	4388		
DB	149 ATCAGGTGGCCCTGAGGCTGACCCGAGCCAGTGCCTGGTGTCCACGGGG	208		
QY	4389 AGGGCCAAAGGCTGCTGAGCTGATTCTCCAGCTCTGCCCCAGCCTTTCGGCTTCACA	4448		

ORIGIN

	Query Match	8.2%;	Score 387;	DB 14;	Length 543;	
	Best Local Similarity	99.6%;	Pred. NO. 0;			
	Matches 487;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	4209	TTTTTTTTTTTATTTGGTTGTTCTGTTTTCATGAGGTTATCGGACCATGGCTCGAGTTCAG	4268			
Db	29	TTTTTTTTTTTATTTGGTTGTTCTGTTTTCATGAGGTTATCGGACCATGGCTCGAGTTCAG	88			
QY	4269	GCACTTCTGTAGAGACTGTTATTTCTGTAAGATGGTTATTTTAACTCTCTCCACCCC	4328			
Db	89	GCACTTCTGTAGAGACTGTTATTTCTGTAAGATGGTTATTTTAACTCTCTCCACCCC	148			
QY	4329	ATCACGGTGCCCTGAGGGCTGACCCCGAGGCGAGTGGAGCTGCCTGGTGTCACGGGGG	4388			
Db	149	ATCACGGTGCCCTGAGGGCTGACCCCGAGGCGAGTGGAGCTGCCTGGTGTCACGGGGG	208			
QY	4389	AGGCGCAAGGCCCTGCTGAGCTGATTTCTCAGCTGTGCTGCCCCAGCCCTTTCGGCCTTCGACA	4448			

sequence tag for this library is GAGTCGGTG. The cell line was provided by Dr James Martin from University of Iowa. TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix TAG_LIB=UI-H-FLO TAG_SEQ=GAGTCGGTG

ORIGIN

Query Match 8.2%; Score 387; DB 14; Length 721;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4209 TTTTATTTTATTTTCTGTTTTCATGAGTTATCGACCATGGCTGAGCTAG 4268
Db 489 TTTTATTTTATTTTCTGTTTTCATGAGTTATCGACCATGGCTGAGCTAG 430
QY 4269 GCATCTTCTGAGGAGACTGTTATTTCTGTAAGATGGTTATTAACCTCTCCACCCC 4328
Db 429 GCATCTTCTGAGGAGACTGTTATTTCTGTAAGATGGTTATTAACCTCTCCACCCC 370
QY 4329 ATCACGGTGGCCCTGAGGCTGACCCGAGGCGCAGTGGAGCTGCCTGTGTCCACGGGG 4388
Db 369 ATCACGGTGGCCCTGAGGCTGACCCGAGGCGCAGTGGAGCTGCCTGTGTCCACGGGG 310
QY 4389 AGGCGCAAGCCTGCTGAGCTGATTTCTCAGCTGCTGCCCGAGCCTTCCGCTTGCACA 4448
Db 309 AGGCGCAAGCCTGCTGAGCTGATTTCTCAGCTGCTGCCCGAGCCTTCCGCTTGCACA 250
QY 4449 GCACAGAGTGGTCAACCCAGGAGCAGCCAGCAGCTGCTCTTCCCTTCTCGGGGG 4508
Db 249 GCACAGAGTGGTCAACCCAGGAGCAGCCAGCAGCTGCTCTTCCCTTCTCGGGGG 190
QY 4509 AAGGAGCTGCTTCTGCTGCTGTAATCTGCTTTCCTTATGAGCCCAACCCGGCAGCTCAGA 4568
Db 189 AAGGAGCTGCTTCTGCTGCTGTAATCTGCTTTCCTTATGAGCCCAACCCGGCAGCTCAGA 130
QY 4569 CTGTGTTGAGCNGCACTGCGAGCTTTTGTCTCTTGTGGGTATTACACAGCAGCGG 4628
Db 129 CTGTGTTGAGCNGCACTGCGAGCTTTTGTCTCTTGTGGGTATTACACAGCAGCGG 70
QY 4629 ACTTGATTTTGATGTTATTTTAAACCACTTAAATAAGAGTCTGTGCTTTAAAAAAA 4688
Db 69 ACTTGATTTTGATGTTATTTTAAACCACTTAAATAAGAGTCTGTGCTTTAAAAAAA 10
QY 4689 AAAAAAAAA 4697
Db 9 AAAAAAAAA 1

RESULT 29

LOCUS AI498183/c 488 bp mRNA linear EST 14-APR-1999
DEFINITION tm90902.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2165426 3', mRNA sequence.

ACCESSION AI498183
VERSION AI498183.1 GI:4390165
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 488)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/STGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 766 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 460.

FEATURES

source

Location/Qualifiers
1...488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2165426"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGAAGTGGGAGCGCCGATAGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 8.2%; Score 386; DB 9; Length 488;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 486; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4205 GCATTTTTTTTTTATTTGTTTGTCTTTCATGAGTTATCGGACCATGGCTGAGC 4264
Db 488 GCATTTTTTTTTTATTTGTTTGTCTTTCATGAGTTATCGGACCATGGCTGAGC 429
QY 4265 TCAGGCATTTCTGTAGGAGACTGTTATTTCTGTAAGATGGTTATTAACCTCTCTCCA 4324
Db 428 TCAGGCATTTCTGTAGGAGACTGTTATTTCTGTAAGATGGTTATTAACCTCTCTCCA 369
QY 4325 CCCCATCAGGTGGCCCTGAGGCTGACCCGAGGCGCAGTGGAGCTGCTGTGTCCACG 4384
Db 368 CCCCATCAGGTGGCCCTGAGGCTGACCCGAGGCGCAGTGGAGCTGCTGTGTCCACG 309
QY 4385 GGGGAGGCGCCAGGCTGCTGAGCTGATTTCTCAGCTGCTGCCCGCAGCTTCCGCTTG 4444
Db 308 GGGGAGGCGCCAGGCTGCTGAGCTGATTTCTCAGCTGCTGCCCGCAGCTTCCGCTTG 249
QY 4445 CACAGCAGAGGTGGTCAACCCAGGAGCAGCCAGGACCTGTCTCTCTTCCGCTTCTG 4504
Db 248 CACAGCAGAGGTGGTCAACCCAGGAGCAGCCAGGACCTGTCTCTCTTCCGCTTCTG 189
QY 4505 GGGGAGGAGGCTGCTGCTGCTCTTAAGTCTGTTTCTTATGGCCCAACCCGGCCACT 4564
Db 188 GGGGAGGAGGCTGCTGCTGCTCTTAAGTCTGTTTCTTATGGCCCAACCCGGCCACT 129
QY 4565 CAGACTTGTGTTGAAGCTGCACTGGCAGCTTTTGTCTCTCTTGGGTATTCAACAGCC 4624
Db 128 CAGACTTGTGTTGAAGCTGCACTGGCAGCTTTTGTCTCTCTTGGGTATTCAACAGCC 69
QY 4625 AGGACTTGATTTTGATGTTATTTTAAACCACTAATAAAGAGTCTGTGCTTAAAAA 4684
Db 68 AGGACTTGATTTTGATGTTATTTTAAACCACTAATAAAGAGTCTGTGCTTAAAAA 9
QY 4685 AAAAAAAAA 4692
Db 8 AAAAAAAAA 1

RESULT 30

LOCUS BU615776/c
DEFINITION UI-H-DF0-bej-m-08-0-UI.s1 NCI_CGAP DF0 Homo sapiens cDNA clone
UI-H-DF0-bej-m-08-0-UI 3', mRNA sequence.

BU615776 706 bp mRNA linear EST 23-SEP-2002
UI-H-DF0-bej-m-08-0-UI.s1 NCI_CGAP DF0 Homo sapiens cDNA clone

```
ACCESSION BU615776
VERSION BU615776.1 GI:23281991
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 706)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             Location/Qualifiers
     source
       1..706
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="UI-H-DF0-bej-m-08-0-UI"
         /tissue_type="Subchondral Bone"
         /dev_stage="Adult"
         /lab_host="DH10B (Life Technologies)"
         /notes="Organ: Bone; Vector: pTV73-Pac (Pharmacia) with a
         modified polylinker; Site 1: EcoR I; Site 2: Not I;
         NCI-CGAP DF0 is a cDNA library containing the following
         tissue(s): Subchondral Bone. The library was constructed
         according to Bonaldo, Lennon and Soares, Genome Research,
         6:791-806, 1996. First strand cDNA synthesis was primed
         with an oligo-dT primer containing a Not I site. Double
         stranded cDNA was ligated to an EcoR I adaptor, digested
         with Not I, and cloned directionally into pTV73-Pac
         vector. The oligonucleotide used to prime the synthesis of
         first-strand cDNA contains a library tag sequence that is
         located between the Not I site and the (dT)18 tail. The
         sequence tag for this library is GTTAAGCGTC.
         TAG TISSUE=Subchondral bone
         TAG_LIB=UI-H-DF0
         TAG_SEQ=GTTAAGCGTC

ORIGIN
Query Match      8.2%; Score 386; DB 13; Length 706;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 486; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4210 TTTTCTCTAGGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTCTCCACCCCA 4329
Db      |
QY 428  CACTTCTCTAGGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTCTCCACCCCA 369
Db      |
QY 4330 TCACGTGCGCTAGGCGTGCACCGGAGCGCAGTGAGCTGCTGGTGTCACCGGGGA 4389
Db      |
QY 368  TCACGTGCGCTAGGCGTGCACCGGAGCGCAGTGAGCTGCTGGTGTCACCGGGGA 309
Db      |
QY 4390 GGSCCAAGCGCTGCTCAGCTGATTCCTCAGCTGCTGCGCCAGCGCTTCGCGCTTCACAG 4449
Db      |
QY 308  GGSCCAAGCGCTGCTCAGCTGATTCCTCAGCTGCTGCGCCAGCGCTTCGCGCTTCACAG 249
Db      |
QY 4450 CACAGAGGTGGTACCCCAAGGACACGCCAGGACCTGCTCCTTCGCGCTTCGCGGGA 4509
Db      |
QY 248  CACAGAGGTGGTACCCCAAGGACACGCCAGGACCTGCTCCTTCGCGCTTCGCGGGA 189
Db      |

QY 4510 AAGGAGCTGCTTCTGTCCTCTGTAACCTGTTTCTTATGCGCCCAACCGGCACTCAGAC 4569
Db      |
QY 188  AGGAGCTGCTTCTGTCCTCTGTAACCTGTTTCTTATGCGCCCAACCGGCACTCAGAC 129
Db      |
QY 4570 TTGTTGAAGCTGCACTGGCAGCTTTTGTCTCTCTTGGGTATTCACACAGCCAGGA 4629
Db      |
QY 128  TTGTTGAAGCTGCACTGGCAGCTTTTGTCTCTCTTGGGTATTCACACAGCCAGGA 69
Db      |
QY 4630 CTTGATTTGATGATTTTAAACACATTAATAAAGAGTCTGTGCTTAAAAA 4689
Db      |
QY 68  CTTGATTTGATGATTTTAAACACATTAATAAAGAGTCTGTGCTTAAAAA 9
Db      |
QY 4690 AAAAAA 4697
Db      |
QY 8  AAAAAA 1
Db      |

RESULT 31
LOCUS BX370444 970 bp mRNA linear EST 08-MAY-2003
DEFINITION BX370444 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK006Y023 5-PRIME, mRNA sequence.
ACCESSION BX370444
VERSION BX370444.1 GI:30437782
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 970)
AUTHORS Li.W.B., Gruber.C., Jesse,J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 880.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG046ZA04_CS04345_1&cluster=880.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG046ZA04_CS04345_1.

FEATURES             Location/Qualifiers
     source
       1..970
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="CS0DK006Y023"
         /cell_type="HELA CELLS COT 25-NORMALIZED"
         /cell_line="HELA"
         /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
         /notes="1st strand cDNA was primed with a NotI-oligo (dT)
         primer. Five prime end enriched, double-strand cDNA was
         digested with Not I and cloned into the Not I and EcoR V
         sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      8.2%; Score 386; DB 13; Length 970;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 AGGTTTGGGAAGGAGATCAGTTGCTGTGATGAGACATTGAATCTCTGAGTACCCAG 510
Db      |
QY 141  AGGTTTGGGAAGGAGATCAGTTGCTGTGATGAGACATTGAATCTCTGAGTACCCAG 200
Db      |
QY 511  AGGAGAAGCTGGTCTCTGTGCAAGAGTATGCTGAACCTGCTGGAGAGACCCGGAATT 570
Db      |
QY 201  AGGAGAAGCTGGTCTCTGTGCAAGAGTATGCTGAACCTGCTGGAGAGACCCGGAATT 260
Db      |
```

QY 571 CACAGAGCAGATGAAGCTCTTACAGAAAGCAGAGCCAGCTGGTGCAGAGAGGACC 630
 Db 261 CACAGAGCAGATGAAGCTCTTACAGAAAGCAGAGCCAGCTGGTGCAGAGAGGACC 320
 QY 631 ACCTGGCGGTGAGCAGCAGCAAGCGCTCTGGCCCGCAGCAAGCTTGAGAGCCTATGCC 690
 Db 321 ACCTGGCGGTGAGCAGCAGCAAGCGCTCTGGCCCGCAGCAAGCTTGAGAGCCTATGCC 380
 QY 691 GTGAGTGCAGCGGCAACCGCTCCCTCAAGAAAGAGGTGTGCAGCGCGCCGGAGG 750
 Db 381 GTGAGTGCAGCGGCAACCGCTCCCTCAAGAAAGAGGTGTGCAGCGCGCCGGAGG 440
 QY 751 AGCAGAGAGCCAGCAGGAGGTGACCTCGCAGCTTCCAGGTGACACTGAATGACATTCAGC 810
 Db 441 AGCAGAGAGCCAGCAGGAGGTGACCTCGCAGCTTCCAGGTGACACTGAATGACATTCAGC 810
 QY 811 TGCAGATGGAACAGCACAATGAGCGC 836
 Db 501 TGCAGATGGAACAGCACAATGAGCGC 526

RESULT 32
 BM452208 1089 bp mRNA linear EST 05-FEB-2002
 LOCUS ACENOCURT 6386282 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5526569
 5', mRNA sequence.
 ACCESSION BM452208
 VERSION BM452208.1 GI:18501248
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1089)
 AUTHORS NIH-MGC.http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12200 Row: 9 Column: 18
 High quality sequence stop: 647.
 Location/Qualifiers
 1..1089
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5526569"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

FEATURES
 source
 Query Match 8.2%; Score 384; DB 12; Length 1089;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 504; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 2460 CAGTGGTGGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 2519
 Db 179 CAGTGGTGGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 238
 QY 2520 CTTTGGCCATCTTGATGCTGAGGTTCTGTTGGTGAGATCAGGTTGTTGGTAAAG 2579

Db 239 CTTTGGCCATCTTGATGCTGAGGTTCTGTTGGTGAGATCAGGTTGTTGGTAAAG 298
 QY 2580 AAAGAAAGGGCTTCTGATGGCTTTCACCAAGCTTACCTGTGGGTTTCAGTCTCGAGAG 2639
 Db 299 AAAGAAAGGGCTTCTGATGGCTTTCACCAAGCTTACCTGTGGGTTTCAGTCTCGAGAG 358
 QY 2640 GCCACCAACAGTTCCTCCATCAGCACTGTCTCATGCAGCAGTGTCTGGGTCCTCCATGTCAG 2699
 Db 359 GCCACCAACAGTTCCTCCATCAGCACTGTCTCATGCAGCAGTGTCTGGGTCCTCCATGTCAG 418
 QY 2700 CTGCTCTTTGGGTTCAAGGGTTTCTGCTTCTGCTGCCACCCACCCACATGTGCAATCC 2759
 Db 419 CTGCTCTTTGGGTTCAAGGGTTTCTGCTTCTGCTGCCACCCACCCACATGTGCAATCC 478
 QY 2760 TCAGATTTGCTCGATTTCTATTTCCTGGCAGCTCCCTGCTGCTGCTTGGGGATTCTACT 2819
 Db 479 TCAGATTTGCTCGATTTCTATTTCCTGGCAGCTCCCTGCTGCTGCTTGGGGATTCTACT 538
 QY 2820 TCTTCTGCTGGG-GCCCATAGCTGTGTCTAACAGGTAGAAATGAATGAATGAAT 2878
 Db 539 TCTTCTGCTGGGAGCCCATAGCTGTGTCTAACAGGTAGAAATGAATGAATGAAT 598
 QY 2879 GACTGGGCCCCAGAAATCCATAAATGGCTGCAGACAGTGTGTTCTGTCGCTGCTGCTAC 2938
 Db 599 GACTGGGCCCCAGAAATCCATAAATGGCTGCAGACAGTGTGTTCTGTCGCTGCTGCTAC 658
 QY 2939 CCCCACTCCAGTACATTAATCTAT 2963
 Db 659 CCCCACTCCAGTACATTAATCTAT 683

BF720458 539 bp mRNA linear EST 03-JAN-2001
 mab54e11.y1 Soares NMEBA branchial_arch.Mus musculus cDNA clone
 IMAGE:3974132 5', mRNA sequence.
 ACCESSION BF720458
 VERSION BF720458.1 GI:12021460
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 539)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/iresources.shtml
 MGI:1474164
 Seq primer: -40RP from Gibco
 High quality sequence stop: 443.
 Location/Qualifiers
 1..539
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:3974132"
 /tissue_type="branchial arches"
 /dev_stage="embryo, 10.5 dpc"
 /lab_host="DH10B (phage resistant)"
 /clone_lib="Soares NMEBA branchial arch"
 /notes="vector: pT730-Pac (pharmacia) with a modified
 polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA

FEATURES
 source

DEFINITION 7q20b06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3698651 3',
mRNA sequence.
ACCESSION BF478077
VERSION BF478077.1 GI:11548904
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 510)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 386.
FEATURES
Location/Qualifiers
1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3698651"
/tissue type="pooled germ cell tumors"
/lab host="DH10B"
/clone lib="NCI CGAP GC6"
/note="Vector: pFR735-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clone10a
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 8.1%; Score 379; DB 10; Length 510;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4210 TTTTCTCTGAGAGACTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACCCCA 4329
DB 482 TTTTCTCTGAGAGACTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACCCCA 4363
QY 4270 CACTTCTCTGAGAGACTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACCCCA 4329
DB 422 CACTTCTCTGAGAGACTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACCCCA 4363
QY 4330 TCACGGTGCCCTGAGGGTGACCCCGAGGCCAGTGAGCTGCCTGGTTCACCGGGGA 4389
DB 362 TCACGGTGCCCTGAGGGTGACCCCGAGGCCAGTGAGCTGCCTGGTTCACCGGGGA 4389
QY 4390 GGGCCAAAGCCCTGCTGAGCTGATTCCTCAGCTGCTGCCCGAGCCCTTCGCGCTTCACAG 4449
DB 302 GGGCCAAAGCCCTGCTGAGCTGATTCCTCAGCTGCTGCCCGAGCCCTTCGCGCTTCACAG 4449
QY 4450 CACAGAGGTGGTTCACCCCGAGGCCAGCCAGGACCTGCTCTCTTGGCCCTTCCTGGGGGA 4509
DB 242 CACAGAGGTGGTTCACCCCGAGGCCAGCCAGGACCTGCTCTCTTGGCCCTTCCTGGGGGA 4509
QY 4510 AAGGAGCTGCTTCTGCTGCTGTAACCTGCTTATGCGCCCAACCCCGGCACTCAGAC 4569

DB 182 AGGAGCTGCTTCTGCTCCTGTAACCTGCTTCTTATGCGCCAGCCGCGCACTCAGAC 123
QY 4570 TTGTTTGAAGCTGCACTGGCAGCTTTTCTCTCCCTTTGGTATTCACACAGCCAGGA 4629
DB 122 TTGTTTGAAGCTGCACTGGCAGCTTTTCTCTCCCTTTGGTATTCACACAGCCAGGA 63
QY 4630 CTTGATTTTGTATGTTATTTAAACACATTAATAAGAGTCTGTTCCTTTAAAAA 4689
DB 62 CTTGATTTTGTATGTTATTTAAACACATTAATAAGAGTCTGTTCCTTTAAAAA 3
QY 4590 A 4590
DB 2 A 2
RESULT 38
BE876516 954 bp mRNA linear EST 20-OCT-2000
LOCUS 60148262F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:388902 5',
mRNA sequence.
ACCESSION BE876516
VERSION BE876516.1 GI:10325292
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM9669 row: k column: 15
High quality sequence stop: 573.
FEATURES
Location/Qualifiers
1..954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:388902"
/tissue type="large cell carcinoma, undifferentiated"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_69"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 8.1%; Score 379; DB 10; Length 954;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3639 GAAAGGCTACAGTGGCGCCACATAATTGCTGACTCAGTCACTTCACTTCTTAATGCC 3698
DB 1 GAAAGGCTACAGTGGCGCCACATAATTGCTGACTCAGTCACTTCACTTCTTAATGCC 60
QY 3699 TGTGAGGAGCTGAGCTGCTGGATCCAGTGTGGTGTAGAGGCCACAGTGAGCAG 3758
DB 61 TGTGAGGAGCTGAGCTGCTGGATCCAGTGTGGTGTAGAGGCCACAGTGAGCAG 120
QY 3759 GTGCCCCCAGCTGGGTTTCCAGCTCAGGAATGTGGCCCCCAGGCAAGGTGACGCTTTG 3818
DB 121 GTGCCCCCAGCTGGGTTTCCAGCTCAGGAATGTGGCCCCCAGGCAAGGTGACGCTTTG 180

QY 3819 CTCACAGCTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCTACCTT 3878
 Db 181 CTCACAGCTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCTACCTT 240
 QY 3879 TTTCCTCTCTTATGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTTTCCA 3938
 Db 241 TTTCCTCTCTTATGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTTTCCA 300
 QY 3939 AAGTAGCCAGGTCCTGGGCGAGATCAGGGAGGTGCTATCCATGAGTGAAGGCCAGT 3998
 Db 301 AAGTAGCCAGGTCCTGGGCGAGATCAGGGAGGTGCTATCCATGAGTGAAGGCCAGT 360
 QY 3999 GTTCTCTCACCTGGGTGG 4017
 Db 361 GTTCTCTCACCTGGGTGG 379

RESULT 39
 AII46364/c
 LOCUS
 DEFINITION qb41906.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:1702714 3', mRNA linear EST 26-OCT-1998
 mRNA sequence.
 ACCESSION AII46364
 VERSION AII46364.1 GI:3674046
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 476)
 AUTHORS NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTCAP), Tumor Gene Index
 JOURNAL Unpublished (1998)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 756 Std Error: 0.00
 Seq primer: -40ml3 fwd. 5T from Amersham
 High quality sequence stop: 452.
 Location/Qualifiers
 1. .476
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1702714"
 /issue_type="glioblastoma (pooled)"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Brn23"
 /note="Organ: Brain; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(GT) primer (5' TTTTACCAATCTGAATGGAGCGCGGCATATCTTTTTTTTTTTTTTTTTTTT T 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

FEATURES

source
 1. .476
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1702714"
 /issue_type="glioblastoma (pooled)"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Brn23"
 /note="Organ: Brain; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(GT) primer (5' TTTTACCAATCTGAATGGAGCGCGGCATATCTTTTTTTTTTTTTTTTTTTT T 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 8.0%; Score 374; DB 9; Length 476;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 474; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4205 GCACCTTTTTTTTATTTATTTGTTGTTTTCATAGGTTATCGACCAATGGGCTGAGC 4264
 Db 476 GCACCTTTTTTTTATTTATTTGTTGTTTTCATAGGTTATCGACCAATGGGCTGAGC 417
 QY 4265 TCAGGCACTTTCTGTAGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTTCCTCCA 4324
 Db 416 TCAGGCACTTTCTGTAGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTTCCTCCA 357
 QY 4325 CCCCATCAGCGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCTGTGTCCACG 4384
 Db 356 CCCCATCAGCGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCTGTGTCCACG 297
 QY 4385 GGGAGGGCCCAAGGCTGCTGAGCTGATTTCCAGCTGCTGCCAGCCCTTTCCGCTTG 4444
 Db 296 GGGAGGGCCCAAGGCTGCTGAGCTGATTTCCAGCTGCTGCCAGCCCTTTCCGCTTG 237
 QY 4445 CACAGCACAGAGGTGGTCAACCCAGGGAGACCCAGGCACCTGCTCTTGGCCCTTCCTG 4504
 Db 236 CACAGCACAGAGGTGGTCAACCCAGGGAGACCCAGGCACCTGCTCTTGGCCCTTCCTG 177
 QY 4505 GGGGAAAGGAGCTGCTTCTGTCCTGTAACTGCTTTCTTATGAGCCCAACCCGGCACT 4564
 Db 176 GGGGAAAGGAGCTGCTTCTGTCCTGTAACTGCTTTCTTATGAGCCCAACCCGGCACT 117
 QY 4565 CAGACTTGTGTAAGCTGCACTGCGACCTTTTTTGTCTCTCTTTGGGTATTCAACAGGCC 4624
 Db 116 CAGACTTGTGTAAGCTGCACTGCGACCTTTTTTGTCTCTCTTTGGGTATTCAACAGGCC 57
 QY 4625 AGGCACTTGATTTTGTATGTTATTTAAACCACTTAAATAAGAGTCTGTGCTTA 4690
 Db 56 AGGCACTTGATTTTGTATGTTATTTAAACCACTTAAATAAGAGTCTGTGCTTA 1

RESULT 40

BQ231672

LOCUS

DEFINITION AGNCOURT 7591747 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6067290 5', mRNA sequence.

ACCESSION

BQ231672

VERSION

BQ231672.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM13346 row: 1 column: 19

High quality sequence stop: 607.

Location/Qualifiers

1. .875

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6067290"

/issue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_92"

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies."

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 8.0%; Score 374; DB 13; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3575 TTTGCTAGCGGAGAGAGGTGCGAGGAGGAGGATGACCTTCCCTTGAAGAG 3634
DB 320 TTTGCTAGCGGAGAGAGGTGCGAGGAGGAGGATGACCTTCCCTTGAAGAG 379
QY 3635 AGGGGAAAGGCTACAGTGGCCCAATATTCCTGACTCACACTTCAGTACCTCTTAA 3694
DB 380 AGGGGAAAGGCTACAGTGGCCCAATATTCCTGACTCACACTTCAGTACCTCTTAA 439
QY 3695 TGCCTGTGAGGAGGAGTGGAGTGGTGGATCCAGTGTGGTGTGAGGAGGAGTGA 3754
DB 440 TGCCTGTGAGGAGGAGTGGAGTGGTGGATCCAGTGTGGTGTGAGGAGGAGTGA 499
QY 3755 GCAGTGGGCCCGAGCTGGGTTTCCAGGTGAGGAAATGGGGCCCGAGGAGGTCAGCC 3814
DB 500 GCAGTGGGCCCGAGCTGGGTTTCCAGGTGAGGAAATGGGGCCCGAGGAGGTCAGCC 559
QY 3815 TTTGCTCAGCTCATCATGCTTAGCTTCAGGCGAGTCTGCAGATGAGTTCCCTA 3874
DB 560 TTTGCTCAGCTCATCATGCTTAGCTTCAGGCGAGTCTGCAGATGAGTTCCCTA 619
QY 3875 CTTTTTCTCTCTCTCATGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTT 3934
DB 620 CTTTTTCTCTCTCTCATGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTT 679
QY 3935 TCCAAAGTAGCCCA 3948
DB 680 TCCAAAGTAGCCCA 693

RESULT 41

LOCUS naf72e11.x1 NCI_CGAP_Brn66 Homo sapiens cDNA clone IMAGE:4169588
DEFINITION 3', mRNA sequence.

ACCESSION BG056649.1 GI:12521336

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 410)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapsb@mail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 396.

Location/Qualifiers

FEATURES

source

1. 410

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4169588"

/tissue_type="glioblastoma with probably TP53 mutation and

without EGFR amplification"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI_CGAP_Brn66"

/notes="Organ: Brain; Vector: pCMV-SPORT6; Site: 1; NofL;

Site: 2; Sali; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.64 Kb. Constructed by Life

Technologies."

ORIGIN

Query Match 7.9%; Score 373; DB 10; Length 410;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2447 AGGATGCTGTACAGTGGTCCCTGGGTGCTGGCTTCCAGTGTCCACCTCTTACCA 2506
DB 38 AGGATGCTGTACAGTGGTCCCTGGGTGCTGGCTTCCAGTGTCCACCTCTTACCA 97
QY 2507 CCCCACTTGGCTCTTTGCCATCTTGATGCTGAGTTTCTGTTGGTGGATCAGTTG 2566
DB 98 CCCCACTTGGCTCTTTGCCATCTTGATGCTGAGTTTCTGTTGGTGGATCAGTTG 157
QY 2567 TTTGTGTAAGAAAGAAAGAGGCTTCTGATGGCTTTGCCAAGCTTACCTGTGGTT 2626
DB 158 TTTGTGTAAGAAAGAGGCTTCTGATGGCTTTGCCAAGCTTACCTGTGGTT 217
QY 2627 TCAGTCTGAGAGGCCACCAACCCAGTTCCTCCATCAGCACTGTCTCCATGACAGTTGCTGG 2686
DB 218 TCAGTCTGAGAGGCCACCAACCCAGTTCCTCCATCAGCACTGTCTCCATGACAGTTGCTGG 277
QY 2687 GTCCCATGTCCAGTGGCTCTTTGGCTTCATGAGTTTCTGCTTCTGCTGCTGCTGCTGCT 2746
DB 278 GTCCCATGTCCAGTGGCTCTTTGGCTTCATGAGTTTCTGCTTCTGCTGCTGCTGCTGCTGCT 337
QY 2747 ACATGTGCAATCTCAAGATTTGCTGATTTCTGCTGATTTCTGCTGCACTTCTGCTGCTGCT 2806
DB 338 ACATGTGCAATCTCAAGATTTGCTGATTTCTGCTGATTTCTGCTGCACTTCTGCTGCTGCT 397
QY 2807 TGGGATTTACT 2819
DB 398 TGGGATTTACT 410

RESULT 42

LOCUS A1419038/c

DEFINITION t53b11.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2102973 3',

mRNA sequence.

ACCESSION A1419038

VERSION A1419038.1 GI:4264969

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 475)

AUTHORS NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute / National Institute of Neurological

Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/STGAP), Tumor Gene Index

Unpublished (1998)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapsb@mail.nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 780 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 442.

Location/Qualifiers

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LL4M1330 row: 1 column: 20
 High quality sequence stop: 546.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5061219"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 70"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 7.9%; Score 370; DB 13; Length 583;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4209 TTTTCTTTTATTTGTTTGTTCATGAGTTATCGACCATGGCTGAGCTCAG 4268
 DB 70 TTTTCTTTTATTTGTTTGTTCATGAGTTATCGACCATGGCTGAGCTCAG 129
 QY 4269 GCACCTTCTGAGGAGCTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACACCC 4328
 DB 130 GCACCTTCTGAGGAGCTGTTATTTCTGTAAGATGTTATTTAACTCCCTCCACACCC 189
 QY 4329 ATCAGGTTGGCTGAGGCTGACCGGAGGCGAGTGAGTGCTGTGTCACAGGGG 4389
 DB 190 ATCAGGTTGGCTGAGGCTGACCGGAGGCGAGTGAGTGCTGTGTCACAGGGG 249
 QY 4389 AGGGCCAGGCTGCTGAGCTGATCTCCAGTGCTGCCCGAGCTTTCCGCTTGCACA 4448
 DB 250 AGGGCCAGGCTGCTGAGCTGATCTCCAGTGCTGCCCGAGCTTTCCGCTTGCACA 309
 QY 4449 GCACAGAGTGTGTCACCCAGGAGACGACGACCTGCTCTTGGCCCTTCCCTGGGG 4508
 DB 310 GCACAGAGTGTGTCACCCAGGAGACGACGACCTGCTCTTGGCCCTTCCCTGGGG 369
 QY 4509 AAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4568
 DB 370 AAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
 QY 4569 CTGTTTGAAGCTGCACCTGGAGCTTTTGTGCTCTTGTGGGTATTCACAGCCAGGG 4628
 DB 430 CTGTTTGAAGCTGCACCTGGAGCTTTTGTGCTCTTGTGGGTATTCACAGCCAGGG 489
 QY 4629 ACTGATTTGATGATTTTAAACCACTAATAAAGAGTCTGTGCTTAA 4680
 DB 490 ACTGATTTGATGATTTTAAACCACTAATAAAGAGTCTGTGCTTAA 541

RESULT 45

BG400724 772 bp mRNA linear EST 12-MAR-2001
 LOCUS 602464108F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4592473 5',
 DEFINITION

mRNA sequence.
 BG400724 GI:13294172
 VERSION BG400724.1 GI:13294172
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LL4M1330 row: 0 column: 02
 High quality sequence stop: 627.
 Location/Qualifiers
 1..772
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4592473"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_75"
 /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctcgcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGACCGGCGGCGCATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.65
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 7.8%; Score 365; DB 12; Length 772;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3653 GGCCCATCAATTCCTGACTTCACACTTCAGTACCTCTTAATGCCCTGTGGAGGACTGG 3712
 DB 1 GGCCCATCAATTCCTGACTTCACACTTCAGTACCTCTTAATGCCCTGTGGAGGACTGG 60
 QY 3713 AGCTGTGGATCCAGTGTGGTGTAGAGGCCACAGTGAGCAGTGCCCGGAGCTGG 3772
 DB 61 AGCTGTGGATCCAGTGTGGTGTAGAGGCCACAGTGAGCAGTGCCCGGAGCTGG 120
 QY 3773 GTTTCCAGGTGAGGAATGTGGCCCGGAGGTCAGGTCAGCTTTGCTCACAGTCCATC 3832
 DB 121 GTTTCCAGGTGAGGAATGTGGCCCGGAGGTCAGGTCAGCTTTGCTCACAGTCCATC 180
 QY 3833 CACTCTAGACCTTCAGGCCAGTCTGCAGATGAGTTCCCTAGCTTTCTTTCTTTCTCAT 3892
 DB 181 CATGCTAGACCTTCAGGCCAGTCTGCAGATGAGTTCCCTAGCTTTCTTTCTTTCTCAT 240
 QY 3893 TGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTTTCCAAAGTAGCCAGGTC 3952
 DB 241 TGACCAATCAACCAATCACTACAGCTGCTCTGCTTCTGCTTTCCAAAGTAGCCAGGTC 300
 QY 3953 CTGGCCAGATGACAGGGGAGGTGCTTATCCATGAGTGAAGCCAGTGTCTTCTCACCTG 4012
 DB 301 CTGGCCAGATGACAGGGGAGGTGCTTATCCATGAGTGAAGCCAGTGTCTTCTCACCTG 360
 QY 4013 GGTGG 4017

[illegible]

QY 4297 GTAAAGATGTTATTAAACCTCTCCACCCATCACGGTGGCCCTGAGGGCTGACCCGG 4356
Db 427 GTAAGATGTTATTAAACCTCTCCACCCATCACGGTGGCCCTGAGGGCTGACCCGG 368
QY 4357 AGCCAGTGGAGCTGCTGCTGCTCCACGGGGAGGGCCCAAGGCTGCTGAGCTGATTTCTC 4416
Db 367 AGCCAGTGGAGCTGCTGCTGCTCCACGGGGAGGGCCCAAGGCTGCTGAGCTGATTTCTC 308
QY 4417 CAGCTGCTGCCAGGCTTTCGCCCTTGCACAGCAGAGGTGGTCCACCCAGGACAGC 4476
Db 307 CAGCTGCTGCCAGGCTTTCGCCCTTGCACAGCAGAGGTGGTCCACCCAGGACAGC 248
QY 4477 CAGGACCTGCTCCCTCTTCCCTTCTGCGGGAAGAGAGCTGCTTCTGCTGCTGTAAC 4536
Db 247 CAGGACCTGCTCCCTCTTCCCTTCTGCGGGAAGAGAGCTGCTTCTGCTGCTGTAAC 188
QY 4537 GCTTCTTATGGCCCAACCGGCGACTCAGACTTGTGTTGAAGCTGCACTGGCAGCTTTT 4596
Db 187 GCTTCTTATGGCCCAACCGGCGACTCAGACTTGTGTTGAAGCTGCACTGGCAGCTTTT 128
QY 4597 TTGCTCTCTTTGGGTATTACAAACAGCAGGAGCTGATTTTGTGTTTAAACCA 4656
Db 127 TTGCTCTCTTTGGGTATTACAAACAGCAGGAGCTGATTTTGTGTTTAAACCA 68
QY 4657 TTAATAAAGAGCTGTTCCCTTAAATAAAAAAAAAAAAAA 4697
Db 67 TTAATAAAGAGCTGTTCCCTTAAATAAAAAAAAAAAAAA 27

RESULT 48
LOCUS A1830542/c
DEFINITION wj51a01.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2406312 3',
mRNA sequence.
ACCESSION A1830542
VERSION A1830542.1 GI:5451213
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 506)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing Arrayed by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/hbrrp/image/image.html
Insert Length: 912 Std Error: 0.00
Seq primer: -40UP from Glbco
High quality sequence stop: 453.
Location/Qualifiers
1. 506
/organism="Homo sapiens"
/ncl_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2406312"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH103 (phage-resistant)"
/clone_lib="NCI_CGAP_Lu19"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from

FEATURES
source

ORIGIN

Query Match 7.6%; Score 359; DB 9; Length 506;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4219 TATTGTTGTTGTTGTTTCCATGAGGTTATCGACCATGAGGCTGAGCTCAGGCACCTTTCTG 4278
Db 461 TATTGTTGTTGTTTTCATGAGGTTATCGACCATGAGGCTGAGCTCAGGCACCTTTCTG 402
QY 4279 TAGGAGACTGTTATTTCTGTAAGAGTGTATTAAACCTCTCCACCCCATCACGGTGG 4338
Db 401 TAGGAGACTGTTATTTCTGTAAGAGTGTATTAAACCTCTCCACCCCATCACGGTGG 342
QY 4339 CCTGAGGGCTGACCCGAGGCGCAGTGGCTGCTGCTCCACGGGGAGGCGCAAGG 4398
Db 341 CCTGAGGGCTGACCCGAGGCGCAGTGGCTGCTGCTCCACGGGGAGGCGCAAGG 282
QY 4399 CCTGCTGAGCTGATTTCTCCAGCTGCTGCCCGAGCCTTTCCGCTTGCACAGCAGAGGT 4458
Db 281 CCTGCTGAGCTGATTTCTCCAGCTGCTGCCCGAGCCTTTCCGCTTGCACAGCAGAGGT 222
QY 4459 GGTCAACCCAGGACAGCAGGACCTGCTCTTCCCTTCTGCGGGAAGAGGCTG 4518
Db 221 GGTCAACCCAGGACAGCAGGACCTGCTCTTCCCTTCTGCGGGAAGAGGCTG 162
QY 4519 CTTTCTGCTCCTGTAACCTGCTTTCTTATGCCCCAACCCGCCACTCAGACTTCTTTGAA 4578
Db 161 CTTTCTGCTCCTGTAACCTGCTTTCTTATGCCCCAACCCGCCACTCAGACTTCTTTGAA 102
QY 4579 GCTGCACTGGCAGCTTTTCTCTCTTTGGGTATTACAAACAGCAGGCTGATTTT 4638
Db 101 GCTGCACTGGCAGCTTTTCTCTCTTTGGGTATTACAAACAGCAGGCTGATTTT 42
QY 4639 GATGTTATTAAACCAATTAATAAAGAGCTGCTTCCCTT 4679
Db 41 GATGTTATTAAACCAATTAATAAAGAGCTGCTTCCCTT 1

RESULT 49
LOCUS CK000671
DEFINITION AGENCOURT 16368557 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30708348 5', mRNA sequence.
ACCESSION CK000671
VERSION CK000671.1 GI:38526705
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 790)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo. "

http://image.llnl.gov
Plate: NDAM1074 row: 9 column: 13
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

source

1..790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708348"
/lab_host="DH10B Tota"
/clone_lib="NIH_MGC 221"
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5kb. Adaptors 5' (AATCGCCAGG)3' and 5'd
(CTCTGCGG)3'. 3' linker sequence - GCGCCGCTGAGGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'd
(ATTACCTCTACTAAGGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCATATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

ORIGIN

Query Match 7.6%; Score 358; DB 14; Length 790;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 638; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 2144 GCGATTTCCTGCTGAAACGAGGACCTGGAGAAATGTTTTCGGTGGGATGTCGG 2203
DB 22 GCGATTTCCTGCTGAAACGAGGACCTGGAGAAATGTTTTCGGTGGGATGTCGG 81
QY 2204 TCAGGAGCCCTTGGGCATCGCTTCCCTGCGCTTTGGTAGTGCCAGGACCGCAATG 2263
DB 82 TCAGGAGCCCTTGGGCATCGCTTCCCTGCGCTTTGGTAGTGCCAGGACCGCAATG 141
QY 2264 ATGCTTCTCAGTAGCCTTATCATTCAGGTGCTCTAGGCTGCAACAAATGATGACA 2323
DB 142 ATGCTTCTCAGTAGCCTTATCATTCAGGTGCTCTAGGCTGCAACAAATGATGACA 201
QY 2324 AGAGATCACCAAGGATTATTCGAAGGTGTTTTTCTTTATTTCTTTTTC--TTTT 2381
DB 202 AGAGATCACCAAGGATTATTCGAAGGTGTTTTTCTTTATTTCTTTTCTTTT 261
QY 2382 TTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2441
DB 262 TTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 321
QY 2442 GGAAGAGGATGCTGACAGTGGTGGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGG 2501
DB 322 GGAAGAGGATGCTGACAGTGGTGGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGG 381
QY 2502 CACACCCACCTTGGCTCTTTGCCATCTTGATGCTGAGGTTTCTTTGTTGGTAGATCA 2561
DB 382 CACACCCACCTTGGCTCTTTGCCATCTTGATGCTGAGGTTTCTTTGTTGGTAGATCA 441
QY 2562 GGTGTTGTTGGTAAAGAAAGAAAGGCTTCTGATGCTTGGCTTGGCAAGCTTACCTGT 2621
DB 442 GGTGTTGTTGGTAAAGAAAGAAAGGCTTCTGATGCTTGGCTTGGCAAGCTTACCTGT 501
QY 2622 GGGTTTCAGTCTGAGAGGCCACCAACAGTTCCTATCAGCACTGTCTCCATGACAGATT 2681
DB 502 GGGTTTCAGTCTGAGAGGCCACCAACAGTTCCTATCAGCACTGTCTCCATGACAGATT 561
QY 2682 GTGGGTCCCATGTCAGTGGCTCTTTGGCTTCATGGGTTTCTGCTTCTGCTGCCCCCA 2741
DB 562 GTGGGTCCCATGTCAGTGGCTCTTTGGCTTCATGGGTTTCTGCTTCTGCTGCCCCCA 621
QY 2742 CCCCCACATGTGCAATCCTCAAGATTGTCTGCTTCTATT 2783

DB 622 CCCCCACATGTGCAATCCTCAAGATTGTCTGCTTCTATT 663

RESULT 50

AL121132
LOCUS DXFzp762H075_r1 762 (synonym: hmcl2) Homo sapiens cDNA clone
DEFINITION DXFzp762H075 5', mRNA sequence.
AL121132
ACCESSION AL121132
VERSION AL121132
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 496)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFzp762H075) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFzp762H075"
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="762 (synonym: hmcl2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 7.6%; Score 356; DB 9; Length 496;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 476; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2576 AAAGAAAGGAAAGGGCTTCTGATGGCTTTCGCCAAGCTTACCTGTGGGTTTCAGTCTG 2635
DB 1 AAAGAAAGGAAAGGGCTTCTGATGGCTTTCGCCAAGCTTACCTGTGGGTTTCAGTCTG 60
QY 2636 AGAGGCCACACACAGTTCCTCATGACAGCTCTCTCCATGAGCAGTTCCTGGTCCCATGT 2695
DB 61 AGAGGCCACACACAGTTCCTCATGACAGCTCTCTCCATGAGCAGTTCCTGGTCCCATGT 120
QY 2696 CCAGTCCCTCTTTGGCTTTCATGGGTTTTCCTGCTTCCCTCCCTCCCTCCCTCCCTCCATGTGA 2755
DB 121 CCAGTCCCTCTTTGGCTTTCATGGGTTTTCCTGCTTCCCTCCCTCCCTCCCTCCATGTGA 180
QY 2756 ATCTCTAAGATTGCTCTGATTCATTTCTGACACCTCCCTGCTGCTTCTGGGATTC 2815
DB 181 ATCTCTAAGATTGCTCTGATTCATTTCTGACACCTCCCTGCTGCTTCTGGGATTC 240
QY 2816 TACTTCTTCTGCTGGG-GCCCATAGCTTCTCTAAAGGTAAGAAATGAATGAAC 2874
DB 241 TACTTCTTCTGCTGGGAGCCCATAGCTTCTCTAAAGGTAAGAAATGAATGAAC 300
QY 2875 TATTGATGGGCCCCAGAAATCATAAATGGTGTGAGACAGTTGTTCTGTCTCTGTT 2934

QY 4297 GTAAGATGGTATTATTAACCTCTCTCCACCCCATCATCGGTGGCCCTGAGGCTGACCCGG 4356
 Db 396 GTAAGATGGTATTATTAACCTCTCTCCACCCCATCATCGGTGGCCCTGAGGCTGACCCGG 337
 QY 4357 AGCCAGTGGAGTGGCTGTGTGTCACCGGGAGGAGCCCAAGGCTGCTGAGTGTATTC 4416
 Db 336 AGCCAGTGGAGTGGCTGTGTGTCACCGGGAGGAGCCCAAGGCTGCTGAGTGTATTC 277
 QY 4417 CAGCTGCTGCCCCAGCCCTTTCCGCTTGCACAGCAGAGGTGGTCAACCCAGGACAGC 4476
 Db 276 CAGCTGCTGCCCCAGCCCTTTCCGCTTGCACAGCAGAGGTGGTCAACCCAGGACAGC 217
 QY 4477 CAGGCACTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4536
 Db 216 CAGGCACTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 157
 QY 4537 GCTTTCTCTATGCGCCCAACCCGCGCCACTCAGACTTGTGTAAGCTGCACTGGCAGCTTT 4596
 Db 156 GCTTTCTCTATGCGCCCAACCCGCGCCACTCAGACTTGTGTAAGCTGCACTGGCAGCTTT 97
 QY 4597 TTGCTCTCTTGGGTATTCACACGCGGAGGACTGATTTGATGATTTTAAACCA 4656
 Db 96 TTGCTCTCTTGGGTATTCACACGCGGAGGACTGATTTGATGATTTTAAACCA 37
 QY 4657 TTAATAAAGAGTCTGTTGCTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 4692
 Db 36 TTAATAAAGAGTCTGTTGCTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 1

RESULT 53

BM053066

LOCUS

DEFINITION ie69g11.y3 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 cDNA clone IMAGE:5672276 5', mRNA sequence.

ACCESSION

BM053066

VERSION

BM053066.1

KEYWORDS

GI:16809013

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 555)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R.,
 Williams, T., Jackson, Y., and Bowers, Y.

TITLE

Endocrine Pancreas Consortium

JOURNAL

Unpublished (2000)

COMMENT

Other ESTs: ie69g11.x2
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@bioh.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@image.llnl.gov
 High quality sequence stop: 429.

FEATURES

source

1..555
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="IMAGE:9606"
 /clone="IMAGE:5672276"
 /sex="Both"
 /tissue_type="Islets of Langerhans"

/dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;
 Site 2: Sal 1; Starting library constructed using
 SuperScript plasmid library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an EcoT of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

ORIGIN

Query Match 7.5%; Score 352; DB 12; Length 555;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3575 TTTCTAGGCGAGAGAGGTGCAGGAAAGAGGAGATATGACCTTCCCTTTGAAGAG 3634
 Db 204 TTTCTAGGCGAGAGAGGTGCAGGAAAGAGGAGATATGACCTTCCCTTTGAAGAG 263
 QY 3635 AGGGAAAGGCTACAGTGGCCACATATGCTGACTCAGCTTACCTTCTTAA 3694
 Db 264 AGGGAAAGGCTACAGTGGCCACATATGCTGACTCAGCTTACCTTCTTAA 323
 QY 3695 TGCCTGTGGAGGAGTCCAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 3754
 Db 324 TGCCTGTGGAGGAGTCCAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 383
 QY 3755 GCAGGTGCCCCAGCTGGGTTTCCAGTCAAGAAATGCGGCCAGGAGTGCAGCC 3814
 Db 384 GCAGGTGCCCCAGCTGGGTTTCCAGTCAAGAAATGCGGCCAGGAGTGCAGCC 443
 QY 3815 TTTCTCAGAGTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCTA 3874
 Db 444 TTTCTCAGAGTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAGGTTCCTA 503
 QY 3875 CCTTTTCTTCTTCTTATGACCAATCAACCAATCACTACAGTGTCTTCG 3926
 Db 504 CCTTTTCTTCTTCTTATGACCAATCAACCAATCACTACAGTGTCTTCG 555

RESULT 54

BE858528/c

LOCUS

DEFINITION BE858528 445 bp mRNA linear EST 29-SEP-2000
 7f96b04.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3304783 3',
 mRNA sequence.

ACCESSION

BE858528

VERSION

BE858528.1

KEYWORDS

GI:10373648

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 445)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute / National Institute of Neurological
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BTGP), Tumor Gene Index
 Unpublished (1998)

JOURNAL

Unpublished (1998)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 389.

FEATURES

source
1. .445
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3304783"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: brain; Vector: pVT3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCATCTGAGTGGAGCGCCGCCGACATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pVT3D vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 7.3%; Score 343; DB 10; Length 445;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 443; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4235 TCCATGAGTTATCGGACCATGGCTGAGCTCAGACACTTCTGTAGGAGACTGTATT 4294
Db 445 TCCATGAGTTATCGGACCATGGCTGAGCTCAGGACCTTCTGTAGGAGACTGTATT 386
QY 4295 CTGTAAAGATGTTATTTAAACCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGACCC 4354
Db 385 CTGTAAAGATGTTATTTAAACCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGACCC 326
QY 4355 GGAGGCCAGTGAGCTGCTGTGTGTCCACGGGGAGGGCCAGGCTGCTGAGCTGATTC 4414
Db 325 GGAGGCCAGTGAGCTGCTGTGTGTCCACGGGGAGGGCCAGGCTGCTGAGCTGATTC 266
QY 4415 TCAGTGTGCTGCCACAGCTTTTCGGCTTGCACAGACAGAGTGGTCAACCCAGGGACA 4474
Db 265 TCAGTGTGCTGCCACAGCTTTTCGGCTTGCACAGACAGAGTGGTCAACCCAGGGACA 206
QY 4475 GCAGGCACTGCTCTTGTGCTTCTTGTGGGGAAAGAGCTGCTTCTGTCCCTGTAA 4534
Db 205 GCAGGCACTGCTCTTGTGCTTCTTGTGGGGAAAGAGCTGCTTCTGTCCCTGTAA 146
QY 4535 CTGCTTCTTATGGCCCAACCGGCACTCAGACTTGTTCAGCTGCACTGGCAGCTT 4594
Db 145 CTGCTTCTTATGGCCCAACCGGCACTCAGACTTGTTCAGCTGCACTGGCAGCTT 86
QY 4595 TTTTGTCTCTTTGGGTATTCAACAGCCAGGAGCTTGTATTTGATGTATTTAAACA 4654
Db 85 TTTTGTCTCTTTGGGTATTCAACAGCCAGGAGCTTGTATTTGATGTATTTAAACA 26
QY 4655 CATTAAATAAGAGTCTGTGCTT 4679
Db 25 CATTAAATAAGAGTCTGTGCTT 1

RESULT 55

CB161278/c
LOCUS
DEFINITION
K-EST0221182 L17N670205n1 Homo sapiens cDNA clone
L17N670205n1-20-P06 5', mRNA sequence.
CB161278
ACCESSION
VERSION
CB161278.1 GI:28147404
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 456)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: F column: 06
High quality sequence stop: 456.

FEATURES

source
1. .456
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L17N670205n1-20-P06"
/sex="F"
/lab_host="Top10F"
/clone_lib="L17N670205n1"
/note="Organ: Liver; Vector: pVT3D-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN

Query Match 7.3%; Score 343; DB 14; Length 456;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 443; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4232 TTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACTTCTGTAGGAGACTGTTA 4291
Db 445 TTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACTTCTGTAGGAGACTGTTA 386
QY 4292 TTTCTGTAAGATGTTATTTAAACCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGA 4351
Db 385 TTTCTGTAAGATGTTATTTAAACCTCTCTCCACCCCATCACGGTGGCCCTGAGGGCTGA 326
QY 4352 CCGGAGGCCAGTGAGCTGCTGTGTCCACGGGGAGGGCCAGGCTGCTGAGCTGA 4411
Db 325 CCGGAGGCCAGTGAGCTGCTGTGTCCACGGGGAGGGCCAGGCTGCTGAGCTGA 266
QY 4412 TTTCTCAGCTGCTGCCCCAGCCTTTCCGCTTGCACAGACAGAGTGGTCAACCCAGGG 4471
Db 265 TTTCTCAGCTGCTGCCCCAGCCTTTCCGCTTGCACAGACAGAGTGGTCAACCCAGGG 206
QY 4472 ACAGCCAGGCACTGCTGCTTGTGCTTCTGTGGGGAAAGAGCTGCTTCTGTCCCTG 4531
Db 205 ACAGCCAGGCACTGCTGCTTGTGCTTCTGTGGGGAAAGAGCTGCTTCTGTCCCTG 146
QY 4532 TAAGTCTTCTTATGCCCCAACCCGCGCACTCAGACTTGTTTGAAGCTGCACTGGCAG 4591
Db 145 TAAGTCTTCTTATGCCCCAACCCGCGCACTCAGACTTGTTTGAAGCTGCACTGGCAG 86
QY 4592 CTTTTTGTCTCTTTGGGTATTCAACAGCCAGGAGCTGTATTTGATGTATTTAAA 4651
Db 85 CTTTTTGTCTCTTTGGGTATTCAACAGCCAGGAGCTGTATTTGATGTATTTAAA 26
QY 4652 CCACATTAATAAGAGTCTGTTC 4676
Db 25 CCACATTAATAAGAGTCTGTTC 1

RESULT 56

BI66824

LOCUS BI686824 729 bp mRNA linear EST 11-OCT-2001
 DEFINITION 60391985F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402204 5',
 mRNA sequence.
 ACCESSION BI686824
 VERSION BI686824.1 GI:116042497
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 729)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM2025 row: 1 column: 21
 High quality sequence stop: 728.

FEATURES

1..729
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5402204"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 90"
 /notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 7.3%; Score 343; DB 12; Length 729;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3129 TCATTGCTCGAGGTCCTTCACTTCATACCTAGCCACCTCTTTAGCCCTTACCTT 3188
 DB 293 TCATTGCTCGAGGTCCTTCACTTCATACCTAGCCACCTCTTTAGCCCTTACCTT 352
 QY 3189 AATCTCTCAGATAAGTGGTTCACAAAGAAATGTTAAGTACTGAATCATGTGACTGAG 3248
 DB 353 AATCTCTCAGATAAGTGGTTCACAAAGAAATGTTAAGTACTGAATCATGTGACTGAG 412
 QY 3249 ACCAGAGATGGCAATGAATGGCACACCACTTCTCCTTCTCGCCCGCCAGGCGATACC 3308
 DB 413 ACCAGAGATGGCAATGAATGGCACACCACTTCTCCTTCTCGCCCGCCAGGCGATACC 472
 QY 3309 ACTGATCTGCATCAGAGTGGCTGCTATTCTCTGGTGTATCCTTACATCTAGGTGCCTT 3368
 DB 473 ACTGATCTGCATCAGAGTGGCTGCTATTCTCTGGTGTATCCTTACATCTAGGTGCCTT 532
 QY 3369 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTCTGT 3428
 DB 533 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTCTGT 592
 QY 3429 GAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCT 3471
 DB 593 GAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCT 635

RESULT 57

LOCUS BI686824 843 bp mRNA linear EST 29-AUG-2001

DEFINITION 603051661P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5201367 5',
 mRNA sequence.
 ACCESSION BI523918
 VERSION BI523918.1 GI:15348710
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 843)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM1504 row: 1 column: 16
 High quality sequence stop: 756.
 Location/Qualifiers
 1..843
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5201367"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /notes="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 7.3%; Score 343; DB 12; Length 843;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3129 TCATTGCTCGAGGTCCTTCACTTCATACCTAGCCACCTCTTTAGCCCTTACCTT 3188
 DB 250 TCATTGCTCGAGGTCCTTCACTTCATACCTAGCCACCTCTTTAGCCCTTACCTT 309
 QY 3189 AATCTCTCAGATAAGTGGTTCACAAAGAAATGTTAAGTACTGAATCATGTGACTGAG 3248
 DB 310 AATCTCTCAGATAAGTGGTTCACAAAGAAATGTTAAGTACTGAATCATGTGACTGAG 369
 QY 3249 ACCAGAGATGGCAATGAATGGCACACCACTTCTCCTTCTCGCCCGCCAGGCGATACC 3308
 DB 370 ACCAGAGATGGCAATGAATGGCACACCACTTCTCCTTCTCGCCCGCCAGGCGATACC 429
 QY 3309 ACTGATCTGCATCAGAGTGGCTGCTATTCTCTGGTGTATCCTTACATCTAGGTGCCTT 3368
 DB 430 ACTGATCTGCATCAGAGTGGCTGCTATTCTCTGGTGTATCCTTACATCTAGGTGCCTT 489
 QY 3369 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTCTGT 3428
 DB 490 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTCTGT 549
 QY 3429 GAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCT 3471
 DB 550 GAAGTGTTCCTCCATGACCTTTTCTTCCCTTTGAATCCCTCT 592

RESULT 58
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BE992847
601433979F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919291 5',
mRNA sequence.
BE992847
BE992847.1 GI:10353429
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 891)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9748 row: m column: 20
High quality sequence stop: 642.
Location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3919291"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES
source

Query Match 7.3%; Score 343; DB 10; Length 891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3129 TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 3188
DB 68 TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 127
QY 3189 AAATCTCTCAGATAAGTTGGTTTCAAGAAGATGTTAAGTACTGAATCATGTGACTGAG 3248
DB 128 AAATCTCTCAGATAAGTTGGTTTCAAGAAGATGTTAAGTACTGAATCATGTGACTGAG 187
QY 3249 ACCAGATGGCAAAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 3308
DB 188 ACCAGATGGCAAAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 247
QY 3309 ACTGATCTGCATCAGAGTTGCCTGCTATTCTCTGGTGATCTCTCACTAGTGGCCCT 3369
DB 248 ACTGATCTGCATCAGAGTTGCCTGCTATTCTCTGGTGATCTCTCACTAGTGGCCCT 307
QY 3369 CAAGCAGCTGTGTGAGTGTGAGATCTGCGCATCTCTGGCTGAGATCTGCTGCTGCT 3428
DB 308 CAAGCAGCTGTGTGAGTGTGAGATCTGCGCATCTCTGGCTGAGATCTGCTGCTGCT 367
QY 3429 GAAGTGTCTCCCATGACCTTTTCTCTCCCTTTGATCCCTCT 3471
DB 368 GAAGTGTCTCCCATGACCTTTTCTCTCCCTTTGATCCCTCT 410

RESULT 59
BG743552

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BG743552
602635590F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780390 5',
mRNA sequence.
BG743552
BG743552.1 GI:14054205
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 922)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10639 row: d column: 23
High quality sequence stop: 858.
Location/Qualifiers
1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4780390"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

Query Match 7.3%; Score 343; DB 12; Length 922;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3129 TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 3188
DB 94 TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 153
QY 3189 AAATCTCTCAGATAAGTTGGTTTCAAGAAGATGTTAAGTACTGAATCATGTGACTGAG 3248
DB 154 AAATCTCTCAGATAAGTTGGTTTCAAGAAGATGTTAAGTACTGAATCATGTGACTGAG 213
QY 3249 ACCAGATGGCAAAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 3308
DB 214 ACCAGATGGCAAAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 273
QY 3309 ACTGATCTGCATCAGAGTTGCCTGCTATTCTCTGGTGATCTCTCACTAGTGGCCCT 3368
DB 274 ACTGATCTGCATCAGAGTTGCCTGCTATTCTCTGGTGATCTCTCACTAGTGGCCCT 333
QY 3369 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGCTGAGATCTGCTGCTGCT 3428
DB 334 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGCTGAGATCTGCTGCTGCT 393
QY 3429 GAAGTGTCTCCCATGACCTTTTCTCTCCCTTTGATCCCTCT 3471
DB 394 GAAGTGTCTCCCATGACCTTTTCTCTCCCTTTGATCCCTCT 436

RESULT 60
BG169290

LOCUS
DEFINITION

BG169290
602321019F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4424068 5',
mRNA sequence.

ACCESSION BG169290
 VERSION BG169290.1 GI:12675993
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 927)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10167 row: f column: 05
 High quality sequence stop: 697.

FEATURES
 Location/Qualifiers
 1..927
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4424068"
 /tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 89"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 7.3%; Score 343; DB 12; Length 927;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3129 TCATTGCTCTGCGAGTCCCTTCATCTTCCATACCTAGCCACCTCTTTAGCCCTTACCTT 3188
 Db 90 TCATTGCTCTGCGAGTCCCTTCATCTTCCATACCTAGCCACCTCTTTAGCCCTTACCTT 149

QY 3189 AAATCTCTCAGATAAGTTGGTTCACAAAGATGTTAAGTACTGAATCATGTGTGACTGAG 3248
 Db 150 AAATCTCTCAGATAAGTTGGTTCACAAAGATGTTAAGTACTGAATCATGTGTGACTGAG 209

QY 3249 ACCAGAGATGGCAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 3308
 Db 210 ACCAGAGATGGCAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 269

QY 3309 ACTGATCTGCATCAGAGTGGCCCTGCTATTTCTCTGGTGTATCCTTCAATCTAGGTGCCCT 3368
 Db 270 ACTGATCTGCATCAGAGTGGCCCTGCTATTTCTCTGGTGTATCCTTCAATCTAGGTGCCCT 329

QY 3369 CAAGAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTGT 3428
 Db 330 CAAGAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTGT 389

QY 3429 GAAGTGTTCCTCATGACCTTTTCTTCCCTTTGAATCCCTCT 3471
 Db 390 GAAGTGTTCCTCATGACCTTTTCTTCCCTTTGAATCCCTCT 432

RESULT 61
 BM459633
 LOCUS
 DEFINITION AGENCOURT_6417770 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534259
 5', mRNA sequence.
 ACCESSION BM459633

BM459633.1 GI:18508673
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 943)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12220 row: h column: 04
 High quality sequence start: 6
 High quality sequence stop: 660.

FEATURES
 Location/Qualifiers
 1..943
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5534259"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 2.1 kb."

ORIGIN
 Query Match 7.3%; Score 343; DB 12; Length 943;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3129 TCATTGCTCTGCGAGTCCCTTCATCTTCCATACCTAGCCACCTCTTTAGCCCTTACCTT 3188
 Db 75 TCATTGCTCTGCGAGTCCCTTCATCTTCCATACCTAGCCACCTCTTTAGCCCTTACCTT 134

QY 3189 AAATCTCTCAGATAAGTTGGTTCACAAAGATGTTAAGTACTGAATCATGTGTGACTGAG 3248
 Db 135 AAATCTCTCAGATAAGTTGGTTCACAAAGATGTTAAGTACTGAATCATGTGTGACTGAG 194

QY 3249 ACCAGAGATGGCAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 3308
 Db 195 ACCAGAGATGGCAATGAATGGCACACCATTTCTCTCTCTGCCCCAGGCGAGTACC 254

QY 3309 ACTGATCTGCATCAGAGTGGCCCTGCTATTTCTCTGGTGTATCCTTCAATCTAGGTGCCCT 3368
 Db 255 ACTGATCTGCATCAGAGTGGCCCTGCTATTTCTCTGGTGTATCCTTCAATCTAGGTGCCCT 314

QY 3369 CAAGAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTGT 3428
 Db 315 CAAGAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTGCTGT 374

QY 3429 GAAGTGTTCCTCATGACCTTTTCTTCCCTTTGAATCCCTCT 3471
 Db 375 GAAGTGTTCCTCATGACCTTTTCTTCCCTTTGAATCCCTCT 417

RESULT 62
 BM471221
 LOCUS
 DEFINITION AGENCOURT_6478283 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563119
 5', mRNA sequence.
 ACCESSION BM471221
 VERSION BM471221.1 GI:18520263
 KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 988)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>;
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCFT/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM12293 row: j column: 16
High quality sequence stop: 703.

FEATURES
source
1..988
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5563119"
/tissue_type="melanotic melanoma"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: Skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 7.3%; Score 343; DB 12; Length 988;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3129 TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 3188
Db TCATTGCTCTGAGGTCCTTCATCTTCATACCTAGCCACTCTTTAGCCCTTACCTT 330
3189 AAATCTCTCAGATAGTGGTTCACAAAGATGTTAAGTACTGAATCATGTGACTGAG 3248
Db AAATCTCTCAGATAGTGGTTCACAAAGATGTTAAGTACTGAATCATGTGACTGAG 390
3249 ACCAGATGGCAATGAATGACACACCATTTCTCTCTCTGCCCCAGGCGATACC 3308
Db ACCAGATGGCAATGAATGACACACCATTTCTCTCTCTGCCCCAGGCGATACC 450
3309 ACTGATCTGCATCAGAGTTCCTGCTATTCCTGCTGATATCTTCACATCTAGGTGCCCT 3368
Db ACTGATCTGCATCAGAGTTCCTGCTATTCCTGCTGATATCTTCACATCTAGGTGCCCT 510
3369 CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGCTGAGATATCTGCTGCTCTGT 3428
Db CAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGCTGAGATATCTGCTGCTCTGT 570
3429 GAAGTGTTCCTCCATGACCTTTTCTTCTCCCTTTGAATCCCTCT 3471
Db GAAGTGTTCCTCCATGACCTTTTCTTCTCCCTTTGAATCCCTCT 613

RESULT 63
AI740467/c
LOCUS
DEFINITION w15q04.x1 Soares NSF F8 9W OT PA_P S1 Homo sapiens cDNA clone
IMAGE:2365206 3', mRNA sequence.
ACCESSION AI740467
VERSION AI740467.1 GI:5108755
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 546 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source
1..441
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2365206"
/lab_host="DH10B"
/clone_lib="Soares NSF F8 9W OT PA_P S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHP pool 1:
30384-310919, 323208-325895 Soares NBHP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NBHP pool 1:
758280-760583, 772104-774407 Soares NBHP pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHP
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 7.2%; Score 339; DB 9; Length 441;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 439; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

4247 TCGACCATGGCTGAGCTCAGGCACCTTTCTGTAGGAGACTGTATTCTCTGAAGTAGG 4306
Db TCGACCATGGCTGAGCTCAGGCACCTTTCTGTAGGAGACTGTATTCTCTGAAGTAGG 382
4307 TTATTTAACTCTCTCCACCCCATCAGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGG 4366
Db TTATTTAACTCTCTCCACCCCATCAGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGG 322
4367 AGCTGCTGTGTCTCCAGGGGAGGCCAAGGCTCTGTAGCTGATTTCTCAGCTGTCTC 4426
Db AGCTGCTGTGTCTCCAGGGGAGGCCAAGGCTCTGTAGCTGATTTCTCAGCTGTCTC 262
4427 CCCAGCCTTTCCGCTTTGCACAGCAGAGAGTGTGTACCCAGGAGCAGCCAGCACCTG 4486
Db CCCAGCCTTTCCGCTTTGCACAGCAGAGAGTGTGTACCCAGGAGCAGCCAGCACCTG 202
4487 CTCCTCTTGCCCTTCTCTGGGGAAAGAGCTGCTTCTGTCCCTGTAACTGTTCTCTTA 4546
Db CTCCTCTTGCCCTTCTCTGGGGAAAGAGCTGCTTCTGTCCCTGTAACTGTTCTCTTA 142
4547 TGGCCCAACCCGGCCACTCAGACTTGTGTAAGCTGCACTGGGAGCTTTTGTCTCCTT 4606
Db TGGCCCAACCCGGCCACTCAGACTTGTGTAAGCTGCACTGGGAGCTTTTGTCTCCTT 82
4607 TGGGTATTCAACACAGCAGGAGCTTGTATTTGATGTATTTAAACACATTAATAAG 4666
Db TGGGTATTCAACACAGCAGGAGCTTGTATTTGATGTATTTAAACACATTAATAAG 22
4667 AGTCTGTGCTTTAAAAAAA 4687

```

Db 21 AGTCTGTGCTTAAAAAAA 1
RESULT 64
BO680816 878 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT 8211765 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6258548
DEFINITION 5', mRNA sequence.
ACCESSION BO680816
VERSION BO680816.1 GI:21793495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2414 row: j column: 21
High quality sequence stop: 691.
FEATURES
source
1..878
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6258548"
/tissue_type="melanotic melanoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Query Match 7.2%; Score 339; DB 13; Length 878;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 2615 TACCTGTGGTTTCAGTCTCTGAGAGCCACACAGTTCCTCCATCAGCACTGTCCTCATGC 2674
DB 1 TACCTGTGGTTTCAGTCTCTGAGAGCCACACAGTTCCTCCATCAGCACTGTCCTCATGC 60
QY 2675 AGCAGTTGCTGGGTCCCATGCTCCTCTTTGGCTTCATGGTTTTCTGCTTCT 2734
DB 61 AGCAGTTGCTGGGTCCCATGCTCCTCTTTGGCTTCATGGTTTTCTGCTTCT 120
QY 2735 GCCCCACCCACACATGTCGATCTCTCAAGATTTGTCGATTCATTTCTGGCACCTC 2794
DB 121 GCCCCACCCACACATGTCGATCTCTCAAGATTTGTCGATTCATTTCTGGCACCTC 180
QY 2795 CTGCTGCTGCTTGGGGATTCCTCTCTTCTGCTGGG-GCCCATAGCTGTTGCTAAC 2853
DB 181 CTGCTGCTGCTTGGGGATTCCTCTCTTCTGCTGGGAGCCCATAGCTGTTGCTAAC 240
QY 2854 AGTAAAGAAATGAATGTAATTCGATCGGCCCCAGAAATCCATAAAATGGCTGCAGA 2913
DB 241 AGTAAAGAAATGAATGTAATTCGATCGGCCCCAGAAATCCATAAAATGGCTGCAGA 300

```


IMAGE:1710648 3', mRNA sequence.
 A1139350
 VERSION A1139350.1 GI:3645322
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 472)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1154 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 437.
 Location/Qualifiers
 1..472
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1710648"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Soares_pregnant uterus NbHPU"
 /notes="Organ: uterus; Vector: p7T3-Pac; Site 1: Not 1;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 AACTGGAAGAATCGCGCGCTTTTCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p7T3 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

Query Match 6.8%; Score 319; DB 9; Length 472;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 469; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4217 TTTATTGTTGTTGTTTCCATGAGGTTATCGGACCATGGGTGAGCTCAGGCACTTTC 4276
 DB 472 TTTATTGTTGTTGTTTCCATGAGGTTATCGGACCATGGGTGAGCTCAGGCACTTTC 413

QY 4277 TGAGGAGCTGTTATCTGTAAAGATGTTATTAACCTCTCCACCCCATCACGCT 4336
 DB 412 TGAGGAGCTGTTATCTGTAAAGATGTTATTAACCTCTCCACCCCATCACGCT 353

QY 4337 GGCCCTGAGGGCTGACCCGGAGGCCAGTGAGCTGCTGGTGTCCACGGGGAGGCGCAA 4396
 DB 352 GGCCCTGAGGGCTGACCCGGAGGCCAGTGAGCTGCTGGTGTCCACGGGGAGGCGCAA 293

QY 4397 GGCTGTGAGCTGATCTCCAGCTGCTGCCAGGCTTCCGCTTGCACAGCAGAG 4456
 DB 292 GGCTGTGAGCTGATCTCCAGCTGCTGCCAGGCTTCCGCTTGCACAGCAGAG 233

QY 4457 GTGGTCAACCCGAGGACAGCAGGACCTGCTCTCTTGGCCCTTCTCGGGGAAAGGAGC 4516
 DB 232 GTGGTCAACCCGAGGACAGCAGGACCTGCTCTCTTGGCCCTTCTCGGGGAAAGGAGC 173

QY 4517 TGCCTTCTGTCCTGTAACTGCTTTCCTATGACCCCAACCCGCGCACTCAGCTTGTG 4576
 DB 172 TGCCTTCTGTCCTGTAACTGCTTTCCTATGACCCCAACCCGCGCACTCAGCTTGTG 113

QY 4577 AAGCTGCATGGGAGCTTTTCTGCTCTTGGGTATTACAAACAGCCAGGCACTTGATT 4636
 DB 112 AAGCTGCATGGGAGCTTTTCTGCTCTTGGGTATTACAAACAGCCAGGCACTTGATT 53

QY 4637 TTGATGTATTTTAAACACATTAAATAAGAGTCTGTGCTTTAAAAA 4688
 DB 52 TTGATGTATTTTAAACACATTAAATAAGAGTCTGTGCTTTAAAAA 1

RESULT 72
 BU429029 397 bp mRNA linear EST 09-SEP-2002
 LOCUS UI-HF-BNO-aei-g-01-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone
 DEFINITION IMAGE:3064128 5', mRNA sequence.
 ACCESSION BU429029
 VERSION BU429029.1 GI:22767504
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 397)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward.
 Location/Qualifiers
 1..397
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3064128"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /lab_host="DH10B (LTI)"
 /clone_lib="NIH MGC 50"
 /note="Vector: p7T3-Pac; Site 1: Not1; Site 2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (3.5-4.4kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

Query Match 6.8%; Score 318; DB 13; Length 397;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2470 TGGTGCTGCGCTCCAGTGTCCACCTCTTACACACCCCACTTGGCTCTTGGCATC 2529
 DB 8 TGGTGCTGCGCTCCAGTGTCCACCTCTTACACACCCCACTTGGCTCTTGGCATC 67

QY 2530 TTGATGCTGAGTTCCTCTGTTTGGTGATCAGGTTGTTGGTAAAGAAAGAAAGG 2589
 DB 68 TTGATGCTGAGTTCCTCTGTTTGGTGATCAGGTTGTTGGTAAAGAAAGAAAGG 127

QY 2590 GCTTCTGATGCTTTGCCAAGCTTACCTGTGGGTTTCAGTCTTGAGAGCCACCA 2649
 DB 128 GCTTCTGATGCTTTGCCAAGCTTACCTGTGGGTTTCAGTCTTGAGAGCCACCA 187

QY 2650 GTTCCCATCAGCACTGCTCCATGCAGTGTGCTGGTCCCATGTCCAGTGCCTCTTT 2709
 DB 188 GTTCCCATCAGCACTGCTCCATGCAGTGTGCTGGTCCCATGTCCAGTGCCTCTTT 247

QY 2710 GGCTTCTAGGGTTTTTCTGCTTCTGCCCCACCCCCACATGTGCAATCTCAAGATTG 2769
 DB 248 GGCTTCTAGGGTTTTTCTGCTTCTGCCCCACCCCCACATGTGCAATCTCAAGATTG 307

ORIGIN
 Query Match 6.8%; Score 318; DB 13; Length 397;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2770 TCGTGAATTCATTCTG 2787
Db 308 TCGTGAATTCATTCTG 325

RESULT 73
CB130906
LOCUS
DEFINITION CB130906 584 bp mRNA linear EST 29-JAN-2003
K-EST018092 L12JSHCO Homo sapiens cDNA clone L12JSHCO-12-G09 5',
mRNA sequence.
ACCESSION CB130906
VERSION CB130906.1 GI:28095350
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 584)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 12 row: G column: 09
High quality sequence stop: 584.

FEATURES
Location/Qualifiers
source
1..584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L12JSHCO-12-G09"
/sex="M"
/cell_line="J-SHC"
/lab_host="Top10F"
/clone_lib="L12JSHCO"
/note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 6.7%; Score 316; DB 14; Length 584;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3156 CCATACCTAGCCCACTCTTTAGCCCTTACCTTAATCTCTCAGATAAGTTGGTTCAAA 3215
Db 1 CCATACCTAGCCCACTCTTTAGCCCTTACCTTAATCTCTCAGATAAGTTGGTTCAAA 60

QY 3216 AGAATGTTAAGTACTGAATCATGTGTGACTGAGACAGAGATGGCAATGAATGCACAC 3275
Db 61 AGAATGTTAAGTACTGAATCATGTGTGACTGAGACAGAGATGGCAATGAATGCACAC 120

QY 3276 CATTCTCTCTCTCTGCCCCAGGAGGTACCATGATCTGCATCAGAGTTGCCTGCTA 3335
Db 121 CATTCTCTCTCTCTGCCCCAGGAGGTACCATGATCTGCATCAGAGTTGCCTGCTA 180

QY 3336 TTCTCTGCTGTATCTCTTACATCATCTAGTGCCCTCAAGCAGCTGTGTGAGTGTGAGATCT 3395
Db 181 TTCTCTGCTGTATCTCTTACATCATCTAGTGCCCTCAAGCAGCTGTGTGAGTGTGAGATCT 240

QY 3396 CTGCCATCTCTGCTGAGATGACTGCTCTCTGTAAGTGTTCCTCCATGACCTTTTCTTC 3455
Db 241 CTGCCATCTCTGCTGAGATGACTGCTCTCTGTAAGTGTTCCTCCATGACCTTTTCTTC 300

QY 3456 CCCTTTGAATCCCTCT 3471
Db 301 CCCTTTGAATCCCTCT 316

RESULT 74
AW812918
LOCUS
DEFINITION RC3-ST0186-250200-018-a05 ST0186 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW812918
VERSION AW812918.1 GI:7905912
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 628)
Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202563
PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=st2-RC3-ST0186-250
200-018-a05&t3=2000-02-25&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 623.
Location/Qualifiers
source
1..628
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0186"
/note="Organ: stomach; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 6.7%; Score 315; DB 10; Length 628;
Best Local Similarity 99.6%; Pred. No. 0;

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1333)
AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing By: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12807 row: d column: 09
 High quality sequence start: 71
 High quality sequence stop: 595.
FEATURES Location/Qualifiers
 1..1333
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5760344"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dr primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
ORIGIN
 Query Match 6.6%; Score 312; DB 12; Length 1333;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3129 TCATTGCTCTGAGTCCCTTCATCTCCATACCTAGCCCTCTTTAGCCCTTACCTT 3188
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 232 TCATTGCTCTGAGTCCCTTCATCTCCATACCTAGCCCTCTTTAGCCCTTACCTT 291
 QY 3189 AAATCTCTCAGATAGTTGGTTTCAAAAGATGTTAAAGTACTGAATCATGTGTGACTGAG 3248
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 292 AAATCTCTCAGATAGTTGGTTTCAAAAGATGTTAAAGTACTGAATCATGTGTGACTGAG 351
 QY 3249 ACCAGATGGCAATGAATGACACACATTTCTCTCTCTCTGCCCCAGGAGTACC 3308
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 352 ACCAGATGGCAATGAATGACACACATTTCTCTCTCTCTGCCCCAGGAGTACC 411
 QY 3309 ACTGATCTGCATCAGATTTGCTGCTATCTCTGCTGATCTCTCACAATCAGTGGCCCT 3368
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 412 ACTGATCTGCATCAGATTTGCTGCTATCTCTGCTGATCTCTCACAATCAGTGGCCCT 471
 QY 3369 CAAGCAGTGTGTGAGTGTGAGATCTCTGCAATCTCTGGGTGAGATCTGCTGCTGT 3428
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 472 CAAGCAGTGTGTGAGTGTGAGATCTCTGCAATCTCTGGGTGAGATCTGCTGCTGT 531
 QY 3429 GAAGTGTTCCTCC 3440
 DB |||||||||||||||
 DB 532 GAAGTGTTCCTCC 543
 RESULT 77
 BF434394

LOCUS BF434394 413 bp mRNA linear EST 29-NOV-2000
DEFINITION 7096111.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE3644253 3',
 mRNA sequence.
ACCESSION BF434394
VERSION BF434394.1 GI:11446682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
 Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
 Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
 1..413
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3644253"
 /tissue_type="fibrotheoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Ov18"
 /note="Organ: ovary; Vector: p773D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAGTGGAGCGCGCGACATTTTATTTT 3'];
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
 Query Match 6.6%; Score 311; DB 10; Length 413;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4238 ATGAGTGTATCGGACCATGGCTGAGCTCAGGCACATTTCTGTAGGAGACTGTTATTTCTG 4297
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 ATGAGTGTATCGGACCATGGCTGAGCTCAGGCACATTTCTGTAGGAGACTGTTATTTCTG 60
 QY 4298 TAAAGATGGTTATTAACCTCTCCACCCCATCAGGTGGCCCTGAGGCTGACCCGGA 4357
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 TAAAGATGGTTATTAACCTCTCCACCCCATCAGGTGGCCCTGAGGCTGACCCGGA 120
 QY 4358 GGCCAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4417
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 GGCCAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 4418 AGCTGTGCCCCAGCCCTTCCGCTTGCACAGCAGAGAGTGTGACCCCGAGGACAGCC 4477
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 AGCTGTGCCCCAGCCCTTCCGCTTGCACAGCAGAGAGTGTGACCCCGAGGACAGCC 240
 QY 4478 AGGCACCTGCTCTCTTGGCCCTTCCGCTTGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCT 4537
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 AGGCACCTGCTCTCTTGGCCCTTCCGCTTGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCT 300
 QY 4538 CTTTCCCTTATGGCCCAACCCGCGCACTCAGACTTGTGTTGAAGTGCACCTGGCAGCTTTT 4597
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 CTTTCCCTTATGGCCCAACCCGCGCACTCAGACTTGTGTTGAAGTGCACCTGGCAGCTTTT 360

QY 4598 TGCTCTCTTGGTATTATCAACAGCAGCAGGACTTCAATTTTGTATTTTAA 4650
 |||||
 Db 361 TGCTCTCTTGGTATTATCAACAGCAGCAGGACTTCAATTTTGTATTTAA 413

RESULT 78
 BX362957 1195 bp mRNA linear EST 05-MAY-2003
 LOCUS BX362957 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DK006Y023 5-PRIME, mRNA sequence.
 ACCESSION BX362957
 VERSION BX362957
 KEYWORDS EST, BX362957.1 GI:30382673
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1195)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 880.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK006AH12QPl&cluster=880.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DK006AH12QPl.
 FEATURES
 Location/Qualifiers
 1..1195
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK006Y023"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="HELA CELLS COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 6.5%; Score 306; DB 13; Length 1195;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 60
 Db 166 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 225
 QY 61 AGCGGCTCTCTGAGTAGAAGAGAGTCCCGGAGCAGCAGGCTCTCTGAGAGCCGG 120
 Db 226 AGCGGCTCTCTGAGTAGAAGAGAGTCCCGGAGCAGCAGGCTCTCTGAGAGCCGG 285
 QY 121 AGCGGCTCAAGCAGGAGGCTCACTCTGGGCGCCCTCTGATGCTCTGAGGAGCTGA 180
 Db 286 AGCGGCTCAAGCAGGAGGCTCACTCTGGGCGCCCTCTGATGCTCTGAGGAGCTGA 345
 QY 181 GCCGCCAATCGGAGACATCTGAGCACATCTGCTGACACATTAACACGAGGAGCCCG 240
 Db 346 GCCGCCAATCGGAGACATCTGAGCACATCTGCTGACACATTAACACGAGGAGCCCG 405
 QY 241 GCAGAGTGGGACAGGCTGAGCGGCTGACCCCGAGTGCAGAGAGTCCCGGACT 300
 Db 406 GCAGAGTGGGACAGGCTGAGCGGCTGACCCCGAGTGCAGAGAGTCCCGGACT 465

ORIGIN

Query Match 6.5%; Score 306; DB 13; Length 1195;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 60
 Db 166 CAAAAGCAGCCAGGACCAACCGGAGCAGGAGCCAGGAGCCAGGAGCCAGCCAGCC 225
 QY 61 AGCGGCTCTCTGAGTAGAAGAGAGTCCCGGAGCAGCAGGCTCTCTGAGAGCCGG 120
 Db 226 AGCGGCTCTCTGAGTAGAAGAGAGTCCCGGAGCAGCAGGCTCTCTGAGAGCCGG 285
 QY 121 AGCGGCTCAAGCAGGAGGCTCACTCTGGGCGCCCTCTGATGCTCTGAGGAGCTGA 180
 Db 286 AGCGGCTCAAGCAGGAGGCTCACTCTGGGCGCCCTCTGATGCTCTGAGGAGCTGA 345
 QY 181 GCCGCCAATCGGAGACATCTGAGCACATCTGCTGACACATTAACACGAGGAGCCCG 240
 Db 346 GCCGCCAATCGGAGACATCTGAGCACATCTGCTGACACATTAACACGAGGAGCCCG 405
 QY 241 GCAGAGTGGGACAGGCTGAGCGGCTGACCCCGAGTGCAGAGAGTCCCGGACT 300
 Db 406 GCAGAGTGGGACAGGCTGAGCGGCTGACCCCGAGTGCAGAGAGTCCCGGACT 465

QY 301 ATGTGG 306
 |||||
 Db 466 ATGTGG 471

RESULT 79
 AW474425 487 bp mRNA linear EST 24-FEB-2000
 LOCUS xs95g01.xl NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:277424 3',
 DEFINITION mRNA sequence.
 ACCESSION AW474425
 VERSION AW474425.1 GI:7044531
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 487)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LUNL at:
 www.bio.lnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 419.
 FEATURES
 Location/Qualifiers
 1..487
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:277424"
 /tissue_type="serous papillary carcinoma, high grade, 2
 pooled tumors"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Ut4"
 /notes="Organ: uterus; Vector: pCMV-SPORT6; Site:1: SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.48 kb. Life Technologies catalog #:
 11542-016"

ORIGIN

Query Match 6.5%; Score 304; DB 10; Length 487;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4224 GTTCTTTTGTTCATGAGTTATCGGACCATGGCTGAGCTCAGGCATTTCTGTAGGA 4283
 Db 1 GTTCTTTTGTTCATGAGTTATCGGACCATGGCTGAGCTCAGGCATTTCTGTAGGA 60
 QY 4284 GACTGTATTCTTAAGATGGTTATTAACTCTCCACCCCATCAGGTGGCCCTG 4343
 Db 61 GACTGTATTCTTAAGATGGTTATTAACTCTCTCCACCCCATCAGGTGGCCCTG 120
 QY 4344 AGGCTCAGCCGAGGCGCAGTGAGCTGCTGTGTCCAGGGGAGGGCCAGGCTGC 4403
 Db 121 AGGCTCAGCCGAGGCGCAGTGAGCTGCTGTGTCCAGGGGAGGGCCAGGCTGC 180
 QY 4404 TGAGCTGATTCTCAGCTGTGTGCCCGAGCTTTCCGCTTTCACAGCAGAGGTGTCA 4463
 Db 181 TGAGCTGATTCTCAGCTGTGTGCCCGAGCTTTCCGCTTTCACAGCAGAGGTGTCA 240
 QY 4464 CCCAGGAGCAGCCAGCAGCTGCTCTCTTTCCTTGGGGGAAAGAGTGCCTTC 4523
 Db 241 CCCAGGAGCAGCCAGCAGCTGCTCTCTTTCCTTGGGGGAAAGAGTGCCTTC 300

QY 4524 TGTCCCTGTAAGTCTTCTTATGCGCCACACCCGCGCACTCAGACTTGTGTTGAAGTCG 4583
 Db 301 TGTCCCTGTAAGTCTTCTTATGCGCCACACCCGCGCACTCAGACTTGTGTTGAAGTCG 360
 QY 4584 ACTGACAGCTTTTGTCTCTCTTGGGTATTCACAAACAGCGGACCTGATTTGATGT 4643
 Db 361 ACTGACAGCTTTTGTCTCTCTTGGGTATTCACAAACAGCGGACCTGATTTGATGT 420
 QY 4644 ATTTTAAACCACTTAAATAAAGAGCTCTGTGCTTTA 4680
 Db 421 ATTTTAAACCACTTAAATAAAGAGCTCTGTGCTTTA 457

RESULT 80
 AI813503/c
 LOCUS wj64902.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407634 3',
 DEFINITION mRNA sequence.
 ACCESSION AI813503
 VERSION AI813503.1 GI:5424718
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 496)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1122 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 458.
 Location/Qualifiers
 1..496
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2407634"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI CGAP Lu19"
 /note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 p7T3 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

FEATURES
 source
 1..496
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2407634"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI CGAP Lu19"
 /note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 p7T3 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

ORIGIN
 Query Match 6.2%; Score 292; DB 9; Length 496;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 492; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 4190 TGAGAGGTGCACAGACATTTTATTTGTTGTTTCCATGAGGTATCG 4249

Db 496 TGAGAGGTGCACAGACATTTTATTTGTTGTTTCCATGAGGTATCG 437
 QY 4250 GACCATGGCTGAGCTCAGGCACCTTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTA 4309
 Db 436 GACCATGGCTGAGCTCAGGCACCTTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTA 377
 QY 4310 TTTAAACCTCTCCACCCCATCAACCGTGGCCCTGAGGGCTGACCCGAGGCCAGTGGAGC 4369
 Db 376 TTTAAACCTCTCCACCCCATCAACCGTGGCCCTGAGGGCTGACCCGAGGCCAGTGGAGC 317
 QY 4370 TGCTGTGTGTCACGGGGAGGCGCCAGGCTCTGTAGCTGATTTCTCCAGCTGCTGCC 4429
 Db 316 TGCTGTGTGTCACGGGGAGGCGCCAGGCTCTGTAGCTGATTTCTCCAGTGGTGGCCCC 257
 QY 4430 AGCCTTTCCGCTTGTGACACAGACAGAGGTGGTTCACCCAGGGACAGCCAGGCACCTGCTC 4489
 Db 256 AGCCTTTCCGCTTGTGACACAGACAGAGGTGGTTCACCCAGGGACAGCCAGGCACCTGCTC 197
 QY 4490 CTCTTGCCCTTCTGGGGAAAGAGCTGCTTCTGTCCTGTAAGCTTTCCTTATGG 4549
 Db 196 CTCTTGCCCTTCTGGGGAAAGAGCTGCTTCTGTCCTGTAAGCTTTCCTTATGG 137
 QY 4550 CCCAACCCGCGCACTCAGACTTGTGTAAGCTGCACTGGCAGCTTTTGTCTCCTTTGG 4609
 Db 136 CCCAGCCGCGCACTCAGACTTGTGTAAGCTGCACTGGCAGCTTTTGTCTCCTTTGG 77
 QY 4610 GTATTCAACAGCCAGGACCTGATTTGATGTTTAAACACATTTAAATAAAGAGT 4669
 Db 76 GTATTCAACAGCCAGGACCTGATTTGATGTTTAAACACATTTAAATAAAGAGT 17
 QY 4670 CTGTGCTCTTAAAAA 4685
 Db 16 CTGTGCTCTTAAAAA 1

RESULT 81
 AI335651/c
 LOCUS q124h05.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1948569 3',
 DEFINITION mRNA sequence.
 ACCESSION AI335651
 VERSION AI335651.1 GI:4072578
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 457)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 647 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 430.
 Location/Qualifiers
 1..457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1948569"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"

FEATURES
 source
 1..457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1948569"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"

/clone lib="NCI CGAP GC4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 6.2%; Score 291; DB 9; Length 457;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 441; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4236 CCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTCTGTAGGAGACTGTATTTC 4295
 DB 444 CCATGAGGTATCGGACCATGGCTGAGCTCAGGCACTTCTGTAGGAGACTGTATTTC 385

QY 4296 TGTAAGATGGTTATTATTAACCTCTCCACCCCATCACGGTGGCCCTGAGGCTGACCG 4355
 DB 384 TGTAAGATGGTTATTATTAACCTCTCCACCCCATCACGGTGGCCCTGAGGCTGACCG 325

QY 4356 GAGCCAGTGGAGCTGCTGTGTCTCACGGGGAGGCGCAAGCCCTGTGAGCTGATTC 4415
 DB 324 GAGCCAGTGGAGCTGCTGTGTCTCACGGGGAGGCGCAAGCCCTGTGAGCTGATTC 265

QY 4416 CGAGCTGCTGCCCGACCTTTCGCCCTTGACAGCACAGAGTGCTGACCCGAGGACAG 4475
 DB 264 CGAGCTGCTGCCCGACCTTTCGCCCTTGACAGCACAGAGTGCTGACCCGAGGACAG 205

QY 4476 CGAGCCAGCTGCTCTCTCTGCTCTCTCGGGGAGGAGCTGCTCTGCTGCTGTAAC 4535
 DB 204 CGAGCCAGCTGCTCTCTCTGCTCTCTCGGGGAGGAGCTGCTCTGCTGCTGTAAC 145

QY 4536 TGCTTTCTTATGCGCCAAACCGGCCCACTCAGACTTGTGAGCTGCACTGCGAGCTTT 4595
 DB 144 TGCTTTCTTATGCGCCAGCGGCCCACTCAGACTTGTGAGCTGCACTGCGAGCTTT 85

QY 4596 TTGTCTCTCTTGGGTATTACAGAGCAGGAGCTGATTTGATGATTTTAAACAC 4655
 DB 84 TTGTCTCTCTTGGGTATTACAGAGCAGGAGCTGATTTGATGATTTTAAACAC 25

QY 4656 ATTAATAAAGAGCTGTGTCCTT 4679
 DB 24 ATTAATAAAGAGCTGTGTCCTT 1

RESULT 82
 AAL15145
 LOCUS 211010.r1 Soares pregnant uterus_NbHPU Homo sapiens cDNA clone
 DEFINITION IMAGE:501547 5', mRNA sequence.

ACCESSION AAL15145
 VERSION AAL15145.1 GI:1670569
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, P., Thierry-Mieg, J.,
 Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
 and Marra, M.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)

PUBMED

PUBMED

COMMENT

Contact: Wilson RK

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 945 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 491.

FEATURES

source

1..616
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3806535"
 /db_xref="taxon:9606"
 /clone="IMAGE:501547"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Soares_pregnant_uterus_NbHPU"
 /note="Organ: uterus; Vector: pT73-Pac; Site.1: Not I;
 Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 AACTGGAGAAATTCGGCGCCCTTTTATTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

ORIGIN

Query Match 6.2%; Score 291; DB 9; Length 616;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4048 CCAAGATCTGTGTGGTTTCTTAGATTGCTAGCTTTCTCCAGGGGACACACAGGCTG 4107
 DB 199 CCAAGATCTGTGTGGTTTCTTAGATTGCTAGCTTTCTCCAGGGGACACACAGGCTG 258

QY 4108 AAGCTCAAGCGCATGGCTCTGCTAATAGTAAATGTTTTCAGGCGCTTGTCCAGCTGA 4167
 DB 259 AAGCTCAAGCGCATGGCTCTGCTAATAGTAAATGTTTTCAGGCGCTTGTCCAGCTGA 318

QY 4168 GAGCTTCATGTCACACAGATTCGAGAGGTGTCAGAGCAGCTTTTATTTTATTTGTTG 4227
 DB 319 GAGCTTCATGTCACACAGATTCGAGAGGTGTCAGAGCAGCTTTTATTTTATTTGTTG 378

QY 4228 TTGTTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACTTTCTGTAGGAGACT 4287
 DB 379 TTGTTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACTTTCTGTAGGAGACT 438

QY 4288 GTTATTTCTCTAAGATGGTTATTACCTCTCCACCCCATCAGGTTGG 4338
 DB 439 GTTATTTCTCTAAGATGGTTATTACCTCTCCACCCCATCAGGTTGG 489

RESULT 83

BU752065/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

PUBMED

COMMENT

BU752065 403 bp mRNA linear EST 10-OCT-2002
 UI-1-BB0-add-f-01-0-UI.s1 NCI CGAP_P14 Homo sapiens cDNA clone
 UI-1-BB0-add-f-01-0-UI 3', mRNA sequence.

BU752065

BU752065.1 GI:23709249

EST

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 403)

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Steven Brown
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=yes

FEATURES
source

Location/Qualifiers
1..403
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-1-BB0-add-f-01-0-UI"
/dev_stage="Full Term"
/tissue_type="Placenta"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Placenta; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP P14 is a cDNA library containing the following tissue(s): Placenta full term. The library was constructed according to Bonaudo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGGAA.
TAG TISSUE=placenta human full term
TAG LIB=UI-1-BB0
TAG_SEQ=AGGAA"

ORIGIN

Query Match 6.1%; Score 287; DB 13; Length 403;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4308 TATTTAACCTCTCCACCCATCAGGTGGCCCTGAGGGCTGACCCGAGGCCAGTGGGA 4367
DB |||||||
QY 4368 GCTGCTGTGTCCACGGGGAGGCCAAGGCTGTGAGCTGATTCCTCAGCTCTGCTCC 4427
DB |||||||
QY 4428 CCAGCCCTTCCGCTTGACAGACAGAGTGTGTACCCAGGACAGCCAGGACCTGC 4487
DB |||||||
QY 4488 TCCTCTGCTCTCTCGGGGAAAGAGCTGCTCTGCTCTGCTCTGCTCTCTCTCTAT 4547
DB |||||||
QY 4548 GGCCTAACCCGCTCAGCTGATTTGTAAGCTGACCTGAGCTTTTCTCTCTCTCT 4607
DB |||||||
QY 4608 GGGTATTCACAGCCAGGCTGATTTGATGATTTTAAACACCATTAATAAGA 4667
DB |||||||
QY 4668 GTCTGTGCTTTAAAAAATAAAAAA 4696
DB |||||||

RESULT 84

BM766498
LOCUS
DEFINITION K-EST0048415 S6SNU620s1 Homo sapiens cDNA clone S6SNU620s1-7-G11
5' mRNA sequence.
ACCESSION BM766498
VERSION BM766498.1 GI:19096113
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 432)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 7 row: G column: 11
High quality sequence stop: 432.
Location/Qualifiers
1..432
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S6SNU620s1-7-G11"
/sex="F"
/tissue_type="Scattering floating"
/cell_type="Scattering floating"
/cell_line="SNU-620"
/lab_host="Top10F"
/clone_lib="S6SNU620s1"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 - 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F with
electroporation method."

ORIGIN

Query Match 6.1%; Score 287; DB 12; Length 432;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA.
Sequence: 389-410, 5'AT rich#Low_complexity (matched complement)
Seq primer: M3 FORWARD
POLYA=No.

FEATURES

source

Location/Qualifiers
1. .431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-afd-o-22-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUI) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bentso-soares@uicwa.edu
TAG_SEQ=None found"

ORIGIN

Query Match 6.1%; Score 286; DB 14; Length 431;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4310 TTTAACTCTCTCCACCCATCAGGTGGCTGAGGGCTGACCGGAGGCGAGTGGAGC 4369
Db TTTAACTCTCTCCACCCATCAGGTGGCTGAGGGCTGACCGGAGGCGAGTGGAGC 80

QY 4370 TGCTGTGTGTCACGGGGAGGCGCCAAAGCCCTGCTGAGTGTATTCACAGTGTCTGCC 4429
Db TGCTGTGTGTCACGGGGAGGCGCCAAAGCCCTGCTGAGTGTATTCACAGTGTCTGCC 140

QY 4430 AGCTTTTCGCTTGCACAGCAGAGGTGTGTCACCCAGGACGACGACGACCTGCTC 4489
Db AGCTTTTCGCTTGCACAGCAGAGGTGTGTCACCCAGGACGACGACGACCTGCTC 200

QY 4490 CTCCTGCTCTCTGCGGAAAGAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 4549
Db CTCCTGCTCTCTGCGGAAAGAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 260

QY 4550 CCCAACCGGCGCACTCAGACTGTTTGAAGCTGCACTGCGAGCTTTTGTCTCCTTTGG 4609
Db CCCAACCGGCGCACTCAGACTGTTTGAAGCTGCACTGCGAGCTTTTGTCTCCTTTGG 320

QY 4610 GTATTACACAGCCAGGAGCTGATTTTGATGTATTTTAAACCAATTAATAAGAGT 4669
Db GTATTACACAGCCAGGAGCTGATTTTGATGTATTTTAAACCAATTAATAAGAGT 380

QY 4670 CTGTTGCTTTAAAAAATAAAAAA 4697
Db CTGTTGCTTTAAAAAATAAAAAA 408

RESULT 87

AW294086/c
LOCUS
DEFINITION
IMAGE:2726726 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

384 bp mRNA linear EST 16-JAN-2000
UI-H-B12-abg-b-08-0-UI.s1 NCI_CGAP_Sub4 Homo sapiens cDNA clone
IMAGE:2726726 3', mRNA sequence.
AW294086
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M3 Forward
POLYA=yes.

FEATURES

source

Location/Qualifiers
1. 384
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2726726"
/lab_host="NCI CGAP Sub4"
/clone_lib="NCI CGAP Sub4"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NCI CGAP Sub4 library is a subtracted library derived from
the NCI-CGAP Sub2 library which is a subtracted library
derived from the NCI-CGAP Sub1 library, which is a
subtracted library derived from B1.B1 constituents a
mixture of 21 normalized or subtracted NCI-CGAP
libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28,
NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5,
NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11,
NCI CGAP Kid2, NCI CGAP Kid3, NCI CGAP Co8, NCI CGAP CLL1,
NCI CGAP Lm2, NCI CGAP Brn23, NCI CGAP Lm5,
NCI CGAP Lm24, NCI CGAP Lm19, NCI CGAP GC4, NCI CGAP GC6,
NCI CGAP Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI CGAP Kid3 pool 1 : LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clonids 1322376-1323811,
1456008-1456775, 1500552-1502855) NCI CGAP Kid5 pool 1 :
LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonids
1323912-1325831, 1471368-1472903, 1492104-1493255)
NCI CGAP Lm5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE
Clonids 1414920-1417991, 1520904-1522439) NCI CGAP GC4
pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
Clonids 1257096-1258631, 1469064-1470983,
1475592-1476743) NCI CGAP Pr22 pool 1 : LLAM 2457-2459,
2758-2759, 3062-3068 (IMAGE Clonids 98508-98679,
1101192-1101959, 1217928-1220615) NCI CGAP Co10 pool 1 :
LLAM 2644-2653, 2871-2872 (IMAGE Clonids 1057416-1061255,
1144584-1145351) Subtraction was performed as previously
described (Bonaldo, Lennon & Soares (1996): Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.)
TAG_TISSUE=lung
TAG_LIB=NCI CGAP_Lm19
TAG_SEQ=GACAGC"

ORIGIN

Query Match 6.0%; Score 282; DB 10; Length 384;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 382; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4313 AACCTCTCTCACCCCATCAGGTGGCCCTGAGGGCTGACCGGAGGCGAGTGGAGTGC 4372
Db AACCTCTCTCACCCCATCAGGTGGCCCTGAGGGCTGACCGGAGGCGAGTGGAGTGC 325

QY 4373 CTGTTGTCACCGGGAGGCGCAAGGCTGTGTAGTGTATTCACAGTGTCTGCC 4432
Db CTGTTGTCACCGGGAGGCGCAAGGCTGTGTAGTGTATTCACAGTGTCTGCC 265

/clone lib="UI-E-Cql"
/note="Organ: eye; Vector: pW73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-Cql is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pW73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the cCAT18 tail. The sequence tag for this library is GCATTAAGG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match	6.0%;	Score 280;	DB 12;	Length 578;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 330;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
3687	CCTCTTAATGCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGTAGGAGGC	3746		
1	CCTCTTAATGCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGTAGGAGGC	60		
3747	CACAGTGAAGAGGTGGCCCCAGCTGGTTCCTCCAGGTCAGGAATGTGGCCCCCAGGCAAG	3806		
61	CACAGTGAAGAGGTGGCCCCAGCTGGTTCCTCCAGGTCAGGAATGTGGCCCCCAGGCAAG	120		
3807	GTGCAGCCTTTGCTCACAAGTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAG	3866		
121	GTGCAGCCTTTGCTCACAAGTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATGAG	180		
3867	GTTCCTCACTCTTTCTCTCTCTTCAATCAGCAAAATCAACCAATCACTACAGCTGCTCTGC	3926		
181	GTTCCTCACTCTTTCTCTCTCTTCAATGACCAATCAACCAATCACTACAGCTGCTCTGC	240		
3927	TTCTGCTTTCCAAAGTAGCCAGGTCTCTGGGCAGATCAGGGGAGGTGCCTTATCCATGA	3986		
241	TTCTGCTTTCCAAAGTAGCCAGGTCTCTGGGCAGATCAGGGGAGGTGCCTTATCCATGA	300		
3987	GTGAAGGCCAGTGTCTTCTCACTGGGTGG	4017		
301	GTGAAGGCCAGTGTCTTCTCACTGGGTGG	331		

RESULT 91

RESULT 91					
BM546750					
LOCUS	561 bp	mrna	linear	EST 06-MAR-2002	
DEFINITION	K-EST0125026 S13KMS5 Homo sapiens cDNA clone S13KMS5-40-G04 5',				
	mrna sequence.				

TITLE	Kim, Y.S.
JOURNAL	21C Frontier Korean EST Project 2001
COMMENT	Unpublished (2002)
	Contact: Kim YS
	Genome Research Center
	Korea Research Institute of Bioscience
	52 Eoeun-dong Yuseong-gu, Daeyeon 30
	Tel : +82-43-860-4470
	Fax : +82-43-860-4409
	Email: yongsung@mail.kribb.re.kr


```

High quality sequence stop: 404.
Location/Qualifiers
1. 491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3294704"
/tissue_type="adenocarc"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pa
/notes="Organ: pancreas;
Site 2: Not1; Cloned un
Average insert size 1.7
11548-013"

```

ORIGIN

Query Match	5.5%;	Score 260;	DB 9;	Length 491;
Best Local Similarity	99.4%;	Prod. No. 0;		
Matches 360;	Conservative	2;	Indels	0; Gaps 0;
QY	4322	CCACCCCATCACGGTGCCTGAGGCGCTACCCGAGGCCAGTGGAGCTGCCTGGTGCC	4381	
DDb	362	CCACCCCATCACGGTGCCTGAGGCGCTACCCGAGGCCAGTGGAGCTGCCTGGTGCC	303	
QY	4382	ACGGGGGAGGGCCCAAGCGCTGTGTGAGCTGATTCTCCAGCTGTGTGCCCAGAGCTTTCGGCC	4441	
DDb	302	ACGGGGGAGGGCCCAAGCGCTGTGTGAGCTGATTCTCCAGCTGTGTGCCCAGAGCTTTCGGCC	243	
QY	4442	TTGCACAGCACAGAGTGGTGTACCCACAGGGACAGCCAGGCACCTGCTCTCTTCGCCCTTC	4501	

4308	TATTAAACCTCTCTCACCCCATCAGTGGCCCTGAGGGCTCACCCGAGGCCACTGGA	4367
68	TATTAAACCTCTCTCACCCCATCAGTGGCCCTGAGGGCTCACCCGAGGCCACTGGA	127
4368	GCTGCTGGTGTCCACGGGGGAGGSCAAAGGCCTGCTGAGCTGATTTCTCCAGTGTGTC	4427
128	GCTGCTGGTGTCCACGGGGGAGGSCAAAGGCCTGCTGAGCTGATTTCTCCAGTGTGTC	187
4428	CCAGCCTTTCCGCTTGCACACACAGAGGTGTACCCACGGACAGCCAGCACCTTC	4487
188	CCAGCCTTTCCGCTTGCACACACAGAGGTGTACCCACGGACAGCCAGCACCTTC	247
4488	TCCTCTGGCCTTTCCTGGGGGAA	4510
248	TCCTCTGGCCTTTCCTGGGGGAA	270

Accession	AI636360	491 bp	mRNA	linear	EST 16-DEC-1999
LOCUS	tz78C09.x1	NCI_CGAP_Pan1	Homo sapiens	cDNA clone	IMAGE:2294704 3',
DEFINITION	mRNA sequence.				
Version	AI636360				
Keywords	AI636360.1	GI:4687630			
Source	EST.				
Organism	Homo sapiens	(human)			
Reference	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
Authors	1 (bases 1 to 491)				
Title	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
Journal	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
Comment	Unpublished (1997)				
	Contact: Robert Strausberg. Ph.D.				

Email: cgapbs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbr/image/image.html
Insert Length: 2176 Std Error: 0.00
Seq primer: -30UP from Gibco

Seq primer: -40UP from Gibco


```
QY 4388 GAGGCCAAGGCTGCTGAGCTGATTCTCCAGCTGCTGCCCCAGCCTTTCCGCCCTTGAC 4447
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 292 GAGGCCAAGGCTGCTGAGCTGATTCTCCAGCTGCTGCCCCAGCCTTTCCGCCCTTGAC 233
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 4448 AGCACAGAGGTGGTCAACCCAGGACAGCCAGGCACTGCTCTCTTCCCTTCTGCGGG 4507
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 232 AGCACAGAGGTGGTCAACCCAGGACAGCCAGGCACTGCTCTCTTCCCTTCTGCGGG 173
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 4508 GAAAGAGCTGCTTCTGCTCCCTGTAAGTCTTCCCTTATGGCCCAACCCGGCCACTCAG 4567
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 172 GAAGGAGCTGCTTCTGCTCCCTGTAAGTCTTCCCTTATGGCCCAACCCGGCCACTCAG 113
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 4568 ACTTCTTTGAAGCTGCACTGGCAGCTTTTCTCTCTTTGGGTATTCAACAGCCAGG 4627
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 112 ACTTCTTTGAAGCTGCACTGGCAGCTTTTCTCTCTTTGGGTATTCAACAGCCAGG 53
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 4628 GACTTGATTTTGAATTTTAAACCAATTTAAATAAAGAGTCTGTTGCCTT 4679
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 52 GACTTGATTTTGAATTTTAAACCAATTTAAATAAAGAGTCTGTTGCCTT 1
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: June 9, 2004, 00:49:27
Job time : 11465.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 13:16:30 ; Search time 1835.06 Seconds
(without alignments)
11676.867 Million cell updates/sec

Title: US-10-023-523-17

Perfect score: 4697

Sequence: 1 caaaaagcagccaggacaa.....ttaaaaaaaaaaaaaaa 4697

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2995936 seqs, 228098010 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4697	100.0	4697	9	US-09-962-055-17
2	4697	100.0	4697	9	US-09-976-740-17
3	4697	100.0	4697	13	US-10-671-242-17
4	4697	100.0	4697	14	US-10-023-529-17
5	4697	100.0	4697	14	US-10-023-523-17
6	4697	100.0	4697	16	US-10-616-187-17
7	1589	33.8	2223	13	US-10-276-774-784
8	1439	30.6	1638	9	US-09-976-740-46
9	1439	30.6	1638	13	US-10-671-242-46
10	1439	30.6	1638	14	US-10-023-529-46
11	1439	30.6	1638	14	US-10-023-523-46
12	1439	30.6	1638	16	US-10-616-187-46
13	844	18.0	22255	9	US-09-976-740-51
14	844	18.0	22255	13	US-10-671-242-51

15	844	18.0	22255	14	US-10-023-529-51	Sequence 51, Appl
16	844	18.0	22255	14	US-10-023-523-51	Sequence 51, Appl
17	844	18.0	22255	16	US-10-616-187-51	Sequence 51, Appl
18	367	7.8	529	14	US-10-040-739-869	Sequence 869, Appl
19	347	7.4	1926	9	US-09-822-849A-167	Sequence 167, Appl
20	273	5.8	507	10	US-09-918-995-3048	Sequence 3048, Appl
21	268	5.7	481	10	US-09-918-995-28710	Sequence 28710, Appl
22	264	5.6	419	10	US-09-918-995-11546	Sequence 11546, Appl
23	235	5.0	235	9	US-09-920-300A-1038	Sequence 1038, Appl
24	235	5.0	235	14	US-10-033-528-1038	Sequence 1038, Appl
25	235	5.0	235	15	US-10-099-926-1038	Sequence 1038, Appl
26	230	4.9	2134	16	US-10-108-260A-1586	Sequence 1586, Appl
27	225	4.8	220	9	US-09-777-564-1375	Sequence 1375, Appl
28	225	4.8	250	15	US-10-015-219-1375	Sequence 1375, Appl
29	218	4.6	482	9	US-09-833-381-1334	Sequence 1334, Appl
30	166	3.5	441	9	US-09-962-436-33	Sequence 33, Appl
31	165	3.5	209	13	US-10-085-783A-11181	Sequence 11181, Appl
32	165	3.5	209	16	US-10-242-535A-11181	Sequence 11181, Appl
33	153	3.3	236	15	US-10-116-713-467	Sequence 467, Appl
34	152	3.2	457	10	US-09-918-995-37202	Sequence 37202, Appl
35	143	3.0	439	10	US-09-918-995-10263	Sequence 10263, Appl
36	140	3.0	405	9	US-09-764-869-467	Sequence 467, Appl
37	140	3.0	405	15	US-10-091-504-467	Sequence 467, Appl
38	140	3.0	405	16	US-10-227-577-467	Sequence 467, Appl
39	129	2.7	416	10	US-09-918-995-14235	Sequence 14235, Appl
40	127	2.7	547	10	US-09-918-995-11012	Sequence 11012, Appl
41	108	2.3	284	15	US-10-060-036-2444	Sequence 2444, Appl
42	59	1.3	4722	9	US-09-962-055-14	Sequence 14, Appl
43	59	1.3	4722	9	US-09-976-740-14	Sequence 14, Appl
44	59	1.3	4722	13	US-10-671-243-14	Sequence 14, Appl
45	59	1.3	4722	14	US-10-023-529-14	Sequence 14, Appl
46	59	1.3	4722	14	US-10-023-523-14	Sequence 14, Appl
47	59	1.3	4722	16	US-10-616-187-14	Sequence 14, Appl
48	44	0.9	522	13	US-10-027-632-28968	Sequence 28968, Appl
49	44	0.9	522	16	US-10-027-632-28968	Sequence 28968, Appl
50	37	0.8	937	12	US-09-876-143-1454	Sequence 1454, Appl
51	34	0.7	14615	13	US-10-221-714A-430	Sequence 430, Appl
52	30	0.6	222	13	US-10-424-599-102640	Sequence 102640, Appl
53	30	0.6	258	13	US-10-424-599-123984	Sequence 123984, Appl
54	30	0.6	408	10	US-09-918-995-7289	Sequence 7289, Appl
55	30	0.6	415	9	US-09-960-352-2223	Sequence 2223, Appl
56	30	0.6	430	13	US-10-424-599-71397	Sequence 71397, Appl
57	30	0.6	479	13	US-10-424-599-113649	Sequence 113649, Appl
58	30	0.6	2366	15	US-10-198-846-13850	Sequence 13850, Appl
59	30	0.6	2696	10	US-09-764-891-9921	Sequence 9921, Appl
60	30	0.6	2696	10	US-09-764-891-9921	Sequence 9923, Appl
61	30	0.6	2752	13	US-10-027-632-112265	Sequence 112265, Appl
62	30	0.6	2752	16	US-10-027-632-112265	Sequence 112265, Appl
63	30	0.6	3673778	15	US-10-313-841-1	Sequence 1, Appl
64	29	0.6	385	13	US-10-424-599-129561	Sequence 129561, Appl
65	29	0.6	476	13	US-10-027-632-143302	Sequence 143302, Appl
66	29	0.6	476	13	US-10-027-632-143303	Sequence 143303, Appl
67	29	0.6	476	16	US-10-027-632-143302	Sequence 143303, Appl
68	29	0.6	476	16	US-10-027-632-143303	Sequence 143303, Appl
69	29	0.6	634	13	US-10-027-632-261090	Sequence 261090, Appl
70	29	0.6	634	16	US-10-027-632-261090	Sequence 261090, Appl
71	29	0.6	801	13	US-10-027-632-117473	Sequence 117473, Appl
72	29	0.6	801	16	US-10-027-632-117473	Sequence 117473, Appl
73	29	0.6	809	13	US-10-027-632-158939	Sequence 158939, Appl
74	29	0.6	809	13	US-10-027-632-158940	Sequence 158940, Appl
75	29	0.6	809	16	US-10-027-632-158939	Sequence 158939, Appl
76	29	0.6	809	16	US-10-027-632-158940	Sequence 158940, Appl
77	29	0.6	1041	13	US-10-027-632-116643	Sequence 116643, Appl
78	29	0.6	1041	13	US-10-027-632-116644	Sequence 116644, Appl
79	29	0.6	1041	13	US-10-027-632-116645	Sequence 116645, Appl
80	29	0.6	1041	16	US-10-027-632-116643	Sequence 116643, Appl
81	29	0.6	1041	16	US-10-027-632-116644	Sequence 116644, Appl
82	29	0.6	1041	16	US-10-027-632-116645	Sequence 116645, Appl
83	29	0.6	10894	9	US-09-764-869-2111	Sequence 2111, Appl
84	29	0.6	10894	15	US-10-091-504-2111	Sequence 2111, Appl
85	29	0.6	10894	16	US-10-227-577-2111	Sequence 2111, Appl
86	29	0.6	47573	12	US-09-997-722-106	Sequence 106, Appl
87	29	0.6	122923	13	US-10-087-192-268	Sequence 268, Appl

c 88 29 0.6 136328 15 US-10-101-510-127 Sequence 127, App
 c 89 29 0.6 175077 13 US-10-087-192-1168 Sequence 1168, Ap
 c 90 29 0.6 822900 16 US-10-292-798-1393 Sequence 1393, Ap
 c 91 28 0.6 36 16 US-10-418-182-65 Sequence 65, Appl
 c 92 28 0.6 280 13 US-10-424-599-70577 Sequence 70577, A
 c 93 28 0.6 333 12 US-09-876-143-359 Sequence 359, App
 c 94 28 0.6 482 15 US-10-060-036-2320 Sequence 2320, Ap
 c 95 28 0.6 515 15 US-10-198-846-9933 Sequence 9933, Ap
 c 96 28 0.6 598 13 US-10-027-632-221256 Sequence 221256,
 c 97 28 0.6 598 16 US-10-027-632-221256 Sequence 221256,
 c 98 28 0.6 809 13 US-10-027-632-160640 Sequence 160640,
 c 99 28 0.6 809 13 US-10-027-632-160641 Sequence 160641,
 c 100 28 0.6 809 16 US-10-027-632-160640 Sequence 160640,

ALIGNMENTS

RESULT 1
 US-09-962-055-17 : Sequence 17, Application US/09962055
 : Patent No. US20020052033A1
 : GENERAL INFORMATION:
 : APPLICANT: Lees, Ann M. S.
 : Lees, Robert S.
 : Arjona, Anibal A.
 : TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
 : BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
 : TREATING ATHEROSCLEROSIS
 : NUMBER OF SEQUENCES: 42
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson P.C.
 : STREET: 225 Franklin Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02110-2804
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq for Windows Version 2.0
 : CURRENT APPLICATION DATA: US/09/962,055
 : FILING DATE: 24-Sep-2001
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/979,608
 : FILING DATE: 26-Nov-1997
 : APPLICATION NUMBER: US 60/031,930
 : FILING DATE: 27-Nov-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Myers, Louis
 : REGISTRATION NUMBER: 35,965
 : REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 617/542-5070
 : TELEFAX: 617/542-8906
 : INFORMATION FOR SEQ ID NO: 17:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4697 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: Coding Sequence
 : LOCATION: 3...1592
 : SEQUENCE DESCRIPTION: SEQ ID NO: 17:
 US-09-962-055-17

Query Match 100.0%; Score 4697; DB 9; Length 4697;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAACAGCCAGGACAAACCGAAGCAGACGACCCAGGAGAGCCAGAGCGGCCAGCC 60
 DB 1 CAAAAGAGAGCCAGGACAAACCGAAGCAGACGACCCAGGAGAGCCAGAGCGGCCAGCC 60
 QY 61 AGGCGGCTCTCTGCAAGTAAAGCAGAGGTCGCGGACAGCAGCGCTCTCTCGAAGCGG 120
 DB 61 AGGCGGCTCTCTGCAAGTAAAGCAGAGGTCGCGGACAGCAGCGCTCTCTCGAAGCGG 120
 QY 121 AGGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCTGTGATGTCTGTAGAGAGTGA 180
 DB 121 AGGGGCTCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCTGTGATGTCTGTAGAGAGTGA 180
 QY 181 GCCCCAACTGGAAGACATATCTAGACACATATCTGTGTGGACAATAACACAGGGGGCCCCG 240
 DB 181 GCCCCAACTGGAAGACATATCTAGACACATATCTGTGTGGACAATAACACAGGGGGCCCCG 240
 QY 241 GCGAGGATGGGGACACAGGTCAGCCGGCTGAACCCGAAGATGCAGAGAACTCCCGACCT 300
 DB 241 GCGAGGATGGGGACACAGGTCAGCCGGCTGAACCCGAAGATGCAGAGAACTCCCGACCT 300
 QY 301 ATGTGCAAGGAATGGGGAGCCTGAACCAACTCCAGTAGTCTATGGAGAGAAGAACCT 360
 DB 301 ATGTGCAAGGAATGGGGAGCCTGAACCAACTCCAGTAGTCTATGGAGAGAAGAACCT 360
 QY 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTGCGAGACCGAGACC 420
 DB 361 CCAAGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTGCGAGACCGAGACC 420
 QY 421 ATCGAAGCCACAGGACAGAAAGAAAGCAAGGTTTGGGAAAGGAGATCAGCTTCTGA 480
 DB 421 ATCGAAGCCACAGGACAGAAAGAAAGCAAGGTTTGGGAAAGGAGATCAGCTTCTGA 480
 QY 481 TGCAGACATTGAATCTCTCTGAGTACCCACAGAGAGAGCTGGCTGTCTGTGAAGAAGT 540
 DB 481 TGCAGACATTGAATCTCTCTGAGTACCCACAGAGAGAGCTGGCTGTCTGTGAAGAAGT 540
 QY 541 ATGCTGAATCTGTGGAGAGACCGGAAATTCAGAGAGAGATGAAGTCTCTCAGAGAAA 600
 DB 541 ATGCTGAATCTGTGGAGAGACCGGAAATTCAGAGAGAGATGAAGTCTCTCAGAGAAA 600
 QY 601 AGCAGAGCAGCTGTGTCAGAGAGAGACCACTGCGCGGTGAGCACAGCAAGCGCGTCC 660
 DB 601 AGCAGAGCAGCTGTGTCAGAGAGAGACCACTGCGCGGTGAGCACAGCAAGCGCGTCC 660
 QY 661 TGGCCCGCAGCAAGCTTTGAGAGCTATGCGGTGAGCTGTCAGCGGCAACCGCTCCCTCA 720
 DB 661 TGGCCCGCAGCAAGCTTTGAGAGCTATGCGGTGAGCTGTCAGCGGCAACCGCTCCCTCA 720
 QY 721 AGGAAGAGGTGTGTCAGCGGCGCGGAGGAGGAGGAGCCAGAGGAGTCACTCCG 780
 DB 721 AGGAAGAGGTGTGTCAGCGGCGCGGAGGAGGAGGAGCCAGAGGAGTCACTCCG 780
 QY 781 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGTGGAAACAGCAATGAGCGCACT 840
 DB 781 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGTGGAAACAGCAATGAGCGCACT 840
 QY 841 CCAAGCTGCCCAAGAGAAATGAGTGGCTGAGAGGCTCAAGAGCTGATTGACAGT 900
 DB 841 CCAAGCTGCCCAAGAGAAATGAGTGGCTGAGAGGCTCAAGAGCTGATTGACAGT 900
 QY 901 ATGAGCTGCGCGAGGAGCATATCGAACAAGTCTTTCAAAACACAGGAGCTACAAACAGC 960
 DB 901 ATGAGCTGCGCGAGGAGCATATCGAACAAGTCTTTCAAAACACAGGAGCTACAAACAGC 960
 QY 961 TGTGATGTCAGAGTCCAGCGGCCAGGAGATGCTAAAGGAGGAGAGAGCGGACCC 1020
 DB 961 TGTGATGTCAGAGTCCAGCGGCCAGGAGATGCTAAAGGAGGAGAGAGCGGACCC 1020
 QY 1021 AGCGGAGAGAGGATTTTCTCTCTGAAAGAGGAGTGTAGTCTCCAGAGAGTGTGTGAGCTGA 1080
 DB 1021 AGCGGAGAGAGGATTTTCTCTCTGAAAGAGGAGTGTAGTCTCCAGAGAGTGTGTGAGCTGA 1080

QY	1081	TGAAGCAGCAAGAGACCCACCTGAAAGCAACAGCTTGCCTCTATACACAGAGAAGTTTGAGG	1140	Db	2161	ACAAGGACCTGGAGAAATGTTTTGCGTGGATGATGCTGGTGCAGAGCCCTTGGGC	2220
Db	1081	TGAAGCAGCAAGAGACCCACCTGAAAGCAACAGCTTGCCTCTATACACAGAGAAGTTTGAGG	1140	QY	2221	ATCGCTCCCTCGCTTGGTAGTGCAGGACAGGCCCAATGATGCTTCTCAGTAGCCT	2280
QY	1141	AGTTCCAGAACACACTTTCCAAAGCAGCGAGGTATTCCACCATTTCAAGCAGGAGATGG	1200	Db	2221	ATCGCTCCCTCGCTTGGTAGTGCAGGACAGGCCCAATGATGCTTCTCAGTAGCCT	2280
Db	1141	AGTTCCAGAACACACTTTCCAAAGCAGCGAGGTATTCCACCATTTCAAGCAGGAGATGG	1200	QY	2281	TATCATTCACAGGTGCCTCTCTAGCTGCACAAATGATGACAAAGAGATCACCCAAAGGA	2340
QY	1201	AAAAGATGACTAAGAAAGATCAAGAACTGGAGAAAGAAACCCACCTGTACCCGGTCCC	1260	Db	2281	TATCATTCACAGGTGCCTCTCTAGCTGCACAAATGATGACAAAGAGATCACCCAAAGGA	2340
Db	1201	AAAAGATGACTAAGAAAGATCAAGAACTGGAGAAAGAAACCCACCTGTACCCGGTCCC	1260	QY	2341	TTATTTCTGAAGGTGTTTTTTCTTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	2400
QY	1261	GGGAGACGACCAAGAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATTAAG	1320	Db	2341	TTATTTCTGAAGGTGTTTTTTCTTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	2400
Db	1261	GGGAGACGACCAAGAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATTAAG	1320	QY	2401	TTTTTTTGCACATGACAGTGTGTTGATTGAGGACCTTCCAAAGGAAAAGGAGATGTGACC	2460
QY	1321	AACCTGAGGCTCTGACAGTAAATCCAAACGGCTGGAGAAAGTGTGCCGGGCACTGCAGA	1380	Db	2401	TTTTTTTGCACATGACAGTGTGTTGATTGAGGACCTTCCAAAGGAAAAGGAGATGTGACC	2460
Db	1321	AACCTGAGGCTCTGACAGTAAATCCAAACGGCTGGAGAAAGTGTGCCGGGCACTGCAGA	1380	QY	2461	AGTGGTGGCTGGGTGGCTCCAGTGTCCACCTCTTCCACCCACCTTGGGTCC	2520
QY	1381	CAGAGCGCAATGACCTGAAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCC	1440	Db	2461	AGTGGTGGCTGGGTGGCTCCAGTGTCCACCTCTTCCACCCACCTTGGGTCC	2520
Db	1381	CAGAGCGCAATGACCTGAAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCC	1440	QY	2521	TTTTGCCATCTTGATCTGAGGTTTTCTGTTTGGTGAGATCAGGTTGTTTGTGTAAGA	2580
QY	1441	TCACTGACAGTGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1500	Db	2521	TTTTGCCATCTTGATCTGAGGTTTTCTGTTTGGTGAGATCAGGTTGTTTGTGTAAGA	2580
Db	1441	TCACTGACAGTGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1500	QY	2581	AAGAAAGGGCTTCTGATGCTTGGCTTTCACCAAGCTTACCTGCTGGGTTCAGTCTGAGAGG	2640
QY	1501	CCAGGCTCAAGAGCGCTTCTGCTACCGAGGAGCACCGAGCACAGAGCATCAGGCCAGA	1560	Db	2581	AAGAAAGGGCTTCTGATGCTTGGCTTTCACCAAGCTTACCTGCTGGGTTCAGTCTGAGAGG	2640
Db	1501	CCAGGCTCAAGAGCGCTTCTGCTACCGAGGAGCACCGAGCACAGAGCATCAGGCCAGA	1560	QY	2641	CCACCACAGTTCCTCATCAGCAGCTGTCTCCATGACAGAGTTCCTGCTGGGTCCCATGCTCCAGC	2700
QY	1561	CTGGGCTCAAGAGCGCTTCTGCTACCGAGGAGCACCGAGCACAGAGCATCAGGCCAGA	1620	Db	2641	CCACCACAGTTCCTCATCAGCAGCTGTCTCCATGACAGAGTTCCTGCTGGGTCCCATGCTCCAGC	2700
Db	1561	CTGGGCTCAAGAGCGCTTCTGCTACCGAGGAGCACCGAGCACAGAGCATCAGGCCAGA	1620	QY	2701	TGCTCTTTGGCTTCATGGGTTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2760
QY	1621	GAAGGAGGCGGAGCCAGCCAGCCCTGCGCCATTAAGGCTCCCATGCTGAGCAGCCCA	1680	Db	2701	TGCTCTTTGGCTTCATGGGTTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2760
Db	1621	GAAGGAGGCGGAGCCAGCCAGCCCTGCGCCATTAAGGCTCCCATGCTGAGCAGCCCA	1680	QY	2761	CAAGATTGCTCTGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2820
QY	1681	TTGCTGAAGCAGGATGTTCTTGAACCTGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG	1740	Db	2761	CAAGATTGCTCTGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2820
Db	1681	TTGCTGAAGCAGGATGTTCTTGAACCTGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG	1740	QY	2821	CTTCTGCTGGGGCCCATAGCTGCTGCTTAAACAGGTAAAGAAATGAATGAATGAATGAATGA	2880
QY	1741	GTGGGTGAGTTTACGTACATAGGAGGATTTGCAAGGCTTGCAATGATTAATACCTG	1800	Db	2821	CTTCTGCTGGGGCCCATAGCTGCTGCTTAAACAGGTAAAGAAATGAATGAATGAATGAATGA	2880
Db	1741	GTGGGTGAGTTTACGTACATAGGAGGATTTGCAAGGCTTGCAATGATTAATACCTG	1800	QY	2881	CTGGGCCCCAGAAATCCATAAATGGCTGCAGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT	2940
QY	1801	TAAAGTACAGTGGCTTGCAATGGGATGGGGTGTGTACAGATGAAGTCAAGTGGCTTG	1860	Db	2881	CTGGGCCCCAGAAATCCATAAATGGCTGCAGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT	2940
Db	1801	TAAAGTACAGTGGCTTGCAATGGGATGGGGTGTGTACAGATGAAGTCAAGTGGCTTG	1860	QY	2941	CACTCCAGTACATAAATACTATCTGCTGCTGAGGCAATTTCTATATGCTGAAATGTTCTG	3000
QY	1861	TCTGTGAGCTGAAGCTCTTGAGAGGGCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920	Db	2941	CACTCCAGTACATAAATACTATCTGCTGCTGAGGCAATTTCTATATGCTGAAATGTTCTG	3000
Db	1861	TCTGTGAGCTGAAGCTCTTGAGAGGGCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920	QY	3001	CTGTTGCAAACTTCCAGGGTATTAGCCAGTGTGTCAGAGGATTTGCGGGGACAC	3060
QY	1921	GCAGAAAGTGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1980	Db	3001	CTGTTGCAAACTTCCAGGGTATTAGCCAGTGTGTCAGAGGATTTGCGGGGACAC	3060
Db	1921	GCAGAAAGTGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1980	QY	3061	AGAATGCTCAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3120
QY	1981	CAGAGCTCAAAACAGTAATACCAAGGTCTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2040	Db	3061	AGAATGCTCAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3120
Db	1981	CAGAGCTCAAAACAGTAATACCAAGGTCTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2040	QY	3121	TGAATAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3180
QY	2041	TGCTTGGTCAAGCTCAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2100	Db	3121	TGAATAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3180
Db	2041	TGCTTGGTCAAGCTCAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2100	QY	3181	CTTACCTTAATCTCTCAGATAAGTGTGTTTCAAGAAATGTTTAAAGTCAATGATGATGATGATGAT	3240
QY	2101	AAAAGGTGCTTGGTCAAGCTCAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2160	Db	3181	CTTACCTTAATCTCTCAGATAAGTGTGTTTCAAGAAATGTTTAAAGTCAATGATGATGATGATGAT	3240
Db	2101	AAAAGGTGCTTGGTCAAGCTCAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2160	QY	3241	TGACTGAGACCAAGATGGCAATGAATGACCACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3300
QY	2161	ACAAGGACCTGGAGATGTTTTGGTGGGATGATGCTGGTGCAGGAGCCCTTGGGC	2220				

121 AGGGGCTCAAGCCAGACGGCTCAGTCTGGGGCCCTTCTGTATCTCTGAGGAGCTGA 180
181 GCGCCCACTGGAAGACATACCTAGACATATCTGTGTGACAAATACACAGGGGGCCCGG 240
181 GCGCCCACTGGAAGACATACCTAGACATATCTGTGTGACAAATACACAGGGGGCCCGG 240
241 GCGAGATGGGACACAGGCTGAGCGGCTGAACCGAGATCAGAGAGTCCCGGACCT 300
241 GCGAGATGGGACACAGGCTGAGCGGCTGAACCGAGATCAGAGAGTCCCGGACCT 300
301 ATGTGCAAGGAATGGGGAGCTGAAACCAACTCCAGTAGTCTATGGAGAGGAAGCACT 360
361 CCAGGGGGATCCAAACACAGAGAGATCCGGAGATGACAGAGTCCGGAGACCGAGACC 420
361 CCAGGGGGATCCAAACACAGAGAGATCCGGAGATGACAGAGTCCGGAGACCGAGACC 420
421 ATCGAAGGCCACAGGAGAGAAAGCAAGGTTTGGGGAAGAGATCAAGTTGCTGA 480
421 ATCGAAGGCCACAGGAGAGAAAGCAAGGTTTGGGGAAGAGATCAAGTTGCTGA 480
481 TGACAGATGTAATCTCTAGTACCCAGAGAGAGTGGCTGTCTGTGCAAGAGT 540
481 TGACAGATGTAATCTCTAGTACCCAGAGAGAGTGGCTGTCTGTGCAAGAGT 540
541 ATGCTGAACTGTGAGAGGACCCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 600
541 ATGCTGAACTGTGAGAGGACCCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAA 600
601 AGCAGAGCAGCTGTGCAAGAGAGGACCACTGCGGCTGAGCAGACAGAGGCGCTCC 660
601 AGCAGAGCAGCTGTGCAAGAGAGGACCACTGCGGCTGAGCAGACAGAGGCGCTCC 660
661 TGGCCCGCAGAGCTGTGAGAGCTATGCGGTGAGCTGAGCGGCACACACCGCTCCCTCA 720
661 TGGCCCGCAGAGCTGTGAGAGCTATGCGGTGAGCTGAGCGGCACACACCGCTCCCTCA 720
721 AGGAAGAGGTGTGACAGCGGCGCGGAGGAGAGGAGAGGAGGAGGAGGAGGAGGAGG 780
721 AGGAAGAGGTGTGACAGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
781 ACTTCAGGTGACATGATGATTCAGTGTGAGATGGAACAGACAAATGAGCGCACT 840
781 ACTTCAGGTGACATGATGATTCAGTGTGAGATGGAACAGACAAATGAGCGCACT 840
841 CCAAGCTGCGCCAAAGAGACATGAGCTGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
841 CCAAGCTGCGCCAAAGAGACATGAGCTGCTGAGAGGCTCAAGAGCTGATGAGCAGT 900
901 ATGAGCTGCGCGAGGAGCATATCGAACAAGTCTTCAACACAGGACCTTACACAGCAGC 960
901 ATGAGCTGCGCGAGGAGCATATCGAACAAGTCTTCAACACAGGACCTTACACAGCAGC 960
961 TGGTGATGCTCAGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
961 TGGTGATGCTCAGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
1021 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1021 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1081 TGAAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
1081 TGAAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
1141 AGTTCCAGAACACATTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
1141 AGTTCCAGAACACATTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
1201 AAAAGATGACTAAGAGATCAAGAGCTCGAGAGAGAAACCAACATGTACCGGTCCCGGT 1260

1201 AAAAGATGACTAAGAGATCAAGAGCTCGAGAGAGAAACCAACATGTACCGGTCCCGGT 1260
1261 GGGAGAGAGCAACAAGGCCCTGCTTGTAGATGCTGAGGAGAAACACAGTCCGGATTAAG 1320
1261 GGGAGAGAGCAACAAGGCCCTGCTTGTAGATGCTGAGGAGAAACACAGTCCGGATTAAG 1320
1321 AACTGAGGGGCTCAGGTAAATCAACCGCTTGGAGAAAGCTGTGCGGGCACTGCAGA 1380
1321 AACTGAGGGGCTCAGGTAAATCAACCGCTTGGAGAAAGCTGTGCGGGCACTGCAGA 1380
1381 CAGAGCGAATGACTGTAACCAAGAGGATCAGGACCTGAGTGTGCTGAGGAGGCTCC 1440
1381 CAGAGCGAATGACTGTAACCAAGAGGATCAGGACCTGAGTGTGCTGAGGAGGCTCC 1440
1441 TCACCTGAGTGGCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
1441 TCACCTGAGTGGCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
1501 CCAGGCTCAGAGAGAGCGCTTGTACCCAGAGGAGCAGGAGCAGAGAGCAGAGCAGG 1560
1501 CCAGGCTCAGAGAGAGCGCTTGTACCCAGAGGAGCAGGAGCAGAGAGCAGAGCAGG 1560
1561 CTGGGCTCAAGAGCCACCTCCGCGAGGAGGCTAGAGAGGCTGGTGTGGTCAATGCTGG 1620
1561 CTGGGCTCAAGAGCCACCTCCGCGAGGAGGCTAGAGAGGCTGGTGTGGTCAATGCTGG 1620
1621 GAAGGAGCGGAGCGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
1621 GAAGGAGCGGAGCGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
1681 TTGCTGAAGCAGAGATGTTTGTGACCTGGCTGGATCTGGCACTTGGCAATTTGGATTTT 1740
1681 TTGCTGAAGCAGAGATGTTTGTGACCTGGCTGGATCTGGCACTTGGCAATTTGGATTTT 1740
1741 GTGGCTCAGATTTTACGTACATAGGGCAATTTGCAAGGCTTGCATTAATGATTAACCTG 1800
1741 GTGGCTCAGATTTTACGTACATAGGGCAATTTGCAAGGCTTGCATTAATGATTAACCTG 1800
1801 TAAGTGTACAGTGGGCTTGCATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
1801 TAAGTGTACAGTGGGCTTGCATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
1861 TCTGTGAGCTGAAGAGCTTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
1861 TCTGTGAGCTGAAGAGCTTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
1921 GCAGAGTGAATTCAGTGTCTGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1921 GCAGAGTGAATTCAGTGTCTGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1981 CAGAGCTCAAAACAAAGTAATACACAAAGGCTTGTGATTTGATTTCTAGACCTGG 2040
1981 CAGAGCTCAAAACAAAGTAATACACAAAGGCTTGTGATTTGATTTCTAGACCTGG 2040
2041 TGGTTGCTCAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
2041 TGGTTGCTCAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
2101 AAAAGGTGTCCCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2101 AAAAGGTGTCCCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2161 ACAAGGAGCTGAGAGATGTTTTTGGCTGGGAGTATGCTGTGGTCAAGAGGAGGAGGAGG 2220
2161 ACAAGGAGCTGAGAGATGTTTTTGGCTGGGAGTATGCTGTGGTCAAGAGGAGGAGGAGG 2220
2221 ATCGCTTCCCTGCGCTTGGTAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
2221 ATCGCTTCCCTGCGCTTGGTAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
2281 TATCATTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
2281 TATCATTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340

Db 4501 CTTGGGGGAAAGAGCTGCTTCTGCTGCTGTAACCTGCTTCTTATGCTCCCAACCCCGGC 4560
Qy 4561 CACTCAGACTTGTTCGAAGCTGCACCTGGCAGCTTTTTCCTTCCTTTGGGTATTCACAC 4620
Db 4561 CACTCAGACTTGTTCGAAGCTGCACCTGGCAGCTTTTTCCTTCCTTTGGGTATTCACAC 4620
Qy 4621 AGCCAGGACTTGAATTTGATGATTTTAAACACATTAATAAGAGTCTGTTGCTTAA 4680
Db 4621 AGCCAGGACTTGAATTTGATGATTTTAAACACATTAATAAGAGTCTGTTGCTTAA 4680
Qy 4681 AAAAAAAAAAAAAA 4697
Db 4681 AAAAAAAAAAAAAA 4697

RESULT 3
US-10-671-242-17
; Sequence 17, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS

FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/671,242
CURRENT FILING DATE: 2003-09-24
PRIOR APPLICATION NUMBER: US/09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 17
LENGTH: 4697
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(1592)
US-10-671-242-17

Query Match 100.0%; Score 4697; DB 13; Length 4697;
Best Local Similarity 100.0%; Pred No. 0;
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAAAGCAGCCAGGACACCGAGACCGAGGAGCCCGAGGAGCCCGAGGAGCCCGAGCC 60
Db 1 CAAAAGCAGCCAGGACACCGAGACCGAGGAGCCCGAGGAGCCCGAGGAGCCCGAGCC 60
Qy 61 AGCGGCTCCTCAGTAGAGCAGAGGTCCCGGAGCAGCAGGCTCCTCGGAGCCGG 120
Db 61 AGCGGCTCCTCAGTAGAGCAGAGGTCCCGGAGCAGCAGGCTCCTCGGAGCCGG 120
Qy 121 AGGGGCTCAAGCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGA 180
Db 121 AGGGGCTCAAGCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGA 180
Qy 181 GCGGCCAATCGAAGCATACTAGACATATGTTGACACATTAACAGGGGGGCCCG 240
Db 181 GCGGCCAATCGAAGCATACTAGACATATGTTGACACATTAACAGGGGGGCCCG 240
Qy 241 GCGAGGATGGGACAGAGGTGAGCGGTGAACCCGAGATCAGAGAAGTCCCGGACT 300

Db 241 GCGAGGATGGGACAGAGGTGAGCGGTGAACCCGAGATGCGAGAGGTCCCGGACT 300
Qy 301 ATGTGGCAAGGAATGGGAGCTCAACCACTCCAGTAGTCTATGGAGAGGAACCT 360
Db 301 ATGTGGCAAGGAATGGGAGCTCAACCACTCCAGTAGTCTATGGAGAGGAACCT 360
Qy 361 CAAAGGGGATTCAAAACACAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACC 420
Db 361 CAAAGGGGATTCAAAACACAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACC 420
Qy 421 ATCGAAGGCCACAGAGAGAAAAGCCAGAGGTTTGGGGAAGAGATCACGTTGCTGA 480
Db 421 ATCGAAGGCCACAGAGAGAAAAGCCAGAGGTTTGGGGAAGAGATCACGTTGCTGA 480
Qy 481 TGCAGACATTGATTAATCTCTGAGTACCCAGAGGAGAACTGGCTGCTCTGTGCAAGAAGT 540
Db 481 TGCAGACATTGATTAATCTCTGAGTACCCAGAGGAGAACTGGCTGCTCTGTGCAAGAAGT 540
Qy 541 ATGCTGAACCTGCGAGAGCACCAGGAAATTCACAGAAGCAGATGAAGCTCTTACAGAAA 600
Db 541 ATGCTGAACCTGCGAGAGCACCAGGAAATTCACAGAAGCAGATGAAGCTCTTACAGAAA 600
Qy 601 AGCAGAGCCAGCTGTCAGAGAGAAAGACCACTGCGCGGTGAGCAGCAAGAGCCGCTCC 660
Db 601 AGCAGAGCCAGCTGTCAGAGAGAAAGACCACTGCGCGGTGAGCAGCAAGAGCCGCTCC 660
Qy 661 TGGCCCGCAGCAAGCTTTCAGAGCCTATGCGGTGAGCTGCGGCGCAACACCGCTCCCTCA 720
Db 661 TGGCCCGCAGCAAGCTTTCAGAGCCTATGCGGTGAGCTGCGGCGCAACACCGCTCCCTCA 720
Qy 721 AGAAGAGAGGTGTCAGCGGCGCGGAGGAGGAGAGCGCAAGAGGTGACCTCGC 780
Db 721 AGAAGAGAGGTGTCAGCGGCGCGGAGGAGGAGAGCGCAAGAGGTGACCTCGC 780
Qy 781 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACACACAATGAGCGCAACT 840
Db 781 ACTTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAACACACAATGAGCGCAACT 840
Qy 841 CCAAGCTGCGCCAGAGAACATGAGCTGCTGAGGCTCAAGAGCTGATTTGAGCAGT 900
Db 841 CCAAGCTGCGCCAGAGAACATGAGCTGCTGAGGCTCAAGAGCTGATTTGAGCAGT 900
Qy 901 ATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAAACACAGGACCTACAACAGCAGC 960
Db 901 ATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAAACACAGGACCTACAACAGCAGC 960
Qy 961 TGGTGGATGCCAGCTCAGCAGGCCAGAGATGCTAAAGAGGAGAGAGAGCGGACCC 1020
Db 961 TGGTGGATGCCAGCTCAGCAGGCCAGAGATGCTAAAGAGGAGAGAGAGCGGACCC 1020
Qy 1021 AGCGGAGAGAGATTTTCTCTGAAAGAGGAGTCCAGAGTCCAGAGATGTTGAGCTGA 1080
Db 1021 AGCGGAGAGAGATTTTCTCTGAAAGAGGAGTCCAGAGTCCAGAGATGTTGAGCTGA 1080
Qy 1081 TGAAGCAGAGAGAGACCCACCTGAAAGCAACAGCTTGCCTTATACAGAGAGTTTGAAG 1140
Db 1081 TGAAGCAGAGAGAGACCCACCTGAAAGCAACAGCTTGCCTTATACAGAGAGTTTGAAG 1140
Qy 1141 AGTTCCAGAACACACTTTCCAAAAGCAGAGGTATTCCACCATTCACAGCAGAGATGG 1200
Db 1141 AGTTCCAGAACACACTTTCCAAAAGCAGAGGTATTCCACCATTCACAGCAGAGATGG 1200
Qy 1201 AAAAGATGACTTAAGAGATCAAGAGCTGGAGAAAGAAACCACTATGACCGTCCCGGT 1260
Db 1201 AAAAGATGACTTAAGAGATCAAGAGCTGGAGAAAGAAACCACTATGACCGTCCCGGT 1260
Qy 1261 GGGAGAGCAGAACAGCCCTCTTTCAGATGCTGAGGAGAAACAGTCCGGATTAAG 1320
Db 1261 GGGAGAGCAGAACAGCCCTCTTTCAGATGCTGAGGAGAAACAGTCCGGATTAAG 1320
Qy 1321 AACTGGAGGCGCTCAGGCTAAATAATCAACCGCTGGAGAGAGTGTGCGGCGACTGAG 1380

QY	2581	AGGAAAGGCTTCTGATGGCTTTTGCCAAAGCTTACCTGTGGGTTTCACTCTGAGAGG	2640
Db	2581		
QY	2641	CAACCAAGGCTTCTGATGGCTTTTGCCAAAGCTTACCTGTGGGTTTCACTCTGAGAGG	2640
Db	2641		
QY	2701	TGCTCTTTGGCTTCAATGAGTTTCTGCTCTGCTCCCTCCCTCCCTCCCTCCCTCCCT	2760
Db	2701		
QY	2761	CAAGATTTGCTGATTTCTATTTCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	2820
Db	2761		
QY	2821	CTTCTCTGCTGGGCTCCCTAGCTGTTGTTCTAAACAGGTAAGAAATGAAATGAAATGAA	2880
Db	2821		
QY	2881	CTGGGCTCCAGAAATCCATAAATAGCTGTTGTTCTAAACAGGTAAGAAATGAAATGAA	2940
Db	2881		
QY	2941	CACTCCAGTACATACTACTATGTTACTGTTAGAGCCATTTCTATATGCTGAATGTTCTG	3000
Db	2941		
QY	3001	CTGTTGCAAACTTCCAGGCTATTAGCCAGTGTGTTGCAAGCAAGTTTTCGGGACAAAC	3060
Db	3001		
QY	3061	AGAACTACAGCAAGGTAAGTGGTGTAGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
Db	3061		
QY	3121	TGAATAGTCAATGCTCCTGAGGCTCCCTTCACTCTCCATACCTAGCCACTCTTTTAGCC	3180
Db	3121		
QY	3181	CTTACCTTAAATCTCTCAGATAAGTGGTTTCAAAAGAAATGTTAGTACTGAATCATGTG	3240
Db	3181		
QY	3241	TGACTCAGACCAAGATGCAATGAATGGCAACCATTTCTCTCTCTCTCTCTCTCTCTCT	3300
Db	3241		
QY	3301	CAGGTACCACTGATCTGCATCAGATGCTGCTATTTCTCTGGTGTATCCTTCACTCTA	3360
Db	3301		
QY	3361	GGTGCCCTCAAGCAGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGTACTGC	3420
Db	3361		
QY	3421	TGTCCTGTGAAGTGTTCCTATGACCTTTCTTCCCTTTGAATCCCTCTTGTCTGGAG	3480
Db	3421		
QY	3481	TAGTCTGCTGCTCTCTGCTCCAGTAGGCTTTTCTTACCCAGCCCTTGTGCCAGG	3540
Db	3481		
QY	3541	CTAAGCTGGTACAAGAGCTGCCAATCTCAGAGTGTGTTGCTAGCGAGAGAGTGCAGGGA	3600
Db	3541		
QY	3601	AGAGGAGAGGTATGCACTTCCCTTTCAGAGAGGGAAGGCTTACAGTGGGCCCA	3660
Db	3601		

QY	3661	TAATTGCTGACTCAGACTTCTCTTAAATGCTGTGGAGGACTGGAGCTGCTG	3720
Db	3661		
QY	3721	GATCCAGTGTGGTGTAGAGGCCACAGTGTAGAGGTGGGCCCCAGCTGGGTTTCCCA	3780
Db	3721		
QY	3781	GGTCAAGATGTGGGCCCCAGCAAGGTGCAGCCTTTGCTCACAGCTCCATCCATGTCTA	3840
Db	3781		
QY	3841	GACCTTCAGGCCAGTCTGCAGATGAGGTTCCTTACCTTTTCTCTCTTCAATGACCAAA	3900
Db	3841		
QY	3901	TCAACCAATCACTACAGTCTGCTTCTGCTTTTCCAAAGTACCCAGGTCTCTGGGCCA	3960
Db	3901		
QY	3961	GATCAGGGAGGTGCCTATCCATGAGTGAAGGCCAGTGTCTTCTCACCTGGTGTCTC	4020
Db	3961		
QY	4021	CACACTTGTGACCTCAGTTTGTAGGACCCAGATCTGTGTGTGTCTTCTAGATTGCTAGC	4080
Db	4021		
QY	4081	TTTTCTCCAGGGACCAACAGAGTGAAGTCAAGAGCGCATGGCTCTGCTAATAGTAA	4140
Db	4081		
QY	4141	ATTCGTTTTCAGGGCTTGTCCAGCTGAGGCTTCACTGTCCACCAAGATCTCTGAGAGGTGTC	4200
Db	4141		
QY	4201	AGCAGCACTTTTTTTTTTTTATTTTGTGTTTTCATGAGGTATTCGACCAATGGGCT	4260
Db	4201		
QY	4261	GAGCTCAGGCATTTCTGTAGGAGCTGTTATTTCTGTAAAGATGTTATTTAACCCCTCC	4320
Db	4261		
QY	4321	TCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGGAGGCCAGTGGAGCTGCTGTGTGC	4380
Db	4321		
QY	4381	CACGGGGAGGGCCCAAGGCTGTGAGCTGATTTCTCAGCTGTGCTGCCAGCCCTTCCGC	4440
Db	4381		
QY	4441	CTTGCACAGCACAGAGTGTGTACCCCGAGGGACAGCCAGCCCTGCTCTTTGCCCTT	4500
Db	4441		
QY	4501	CCTGGGGAAAGAGCTGCTTCTGCTCTGTAACTGCTTTTCTTATGGGCCCAACCCGCG	4560
Db	4501		
QY	4561	CACCTCAGACTGTTGAGCTGCACTGGCAGCTTTTGTCTCTCTCTCTCTCTCTCTCT	4620
Db	4561		
QY	4621	AGCCAGGAGCTTGAATTTTGAATGATTTTAAACCAATTAATAAGAGTCTGTTGCCCTTA	4680
Db	4621		
QY	4681	AAAAA	4697
Db	4681	AAAAA	4697

1561	Db	CTGGGCTCTAAGAGCCCACTCTCGCAGGCGCTAGAGAGCCTGGTGTGGGTCTATGCTGG	1620
1621	QY	GAAGGAGCGCAGCCAGCCAGCGCCCTGGGCCATAAAAGGCTCCCATGCTGAGCAGCCCA	1680
1621	Db	GAAGGAGCGCAGCCAGCCAGCGCCCTGGGCCATAAAAGGCTCCCATGCTGAGCAGCCCA	1680
1681	QY	TTGCTGAAGCCAGATGTTCTTGACCTGGCTGCATCTGGCACTTGGCACTTGGATTTT	1740
1681	Db	TTGCTGAAGCCAGATGTTCTTGACCTGGCTGCATCTGGCACTTGGCACTTGGATTTT	1740
1741	QY	GTGGGTCAGTTTACGTACATAGGCAATTTTGCAAGGCCCTTGCAATGCAATTTATACCTG	1800
1741	Db	GTGGGTCAGTTTACGTACATAGGCAATTTTGCAAGGCCCTTGCAATGCAATTTATACCTG	1800
1801	QY	TAAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGGTGTACAGATGAAGTCACTGGCTTG	1860
1801	Db	TAAAGTGTACAGTGGGCTTGCAATTTGGGATGGGGGTGTGTACAGATGAAGTCACTGGCTTG	1860
1861	QY	TCGTGAGCTGAAGAGCTTTGAGAGGGCTGTGCATCTGAGCTGCCATCACAGTGAAGTTG	1920
1861	Db	TCGTGAGCTGAAGAGCTTTGAGAGGGCTGTGTATCTGTAAGCTGCCATCACAGTGAAGTTG	1920
1921	QY	GCAGAAAGTGACTTGAGCAATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT	1980
1921	Db	GCAGAAAGTGACTTGAGCAATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCTTT	1980
1981	QY	CAGAGCTCAAAACAGTAATACACCAAGTCTTTGACGTGCATTTGTCTGTGACAGGGCT	2040
1981	Db	CAGAGCTCAAAACAGTAATACACCAAGTCTTTGACGTGCATTTGTCTGTGACAGGGCT	2040
2041	QY	TGCTTGCTCAGCTCAGGCCCTCCTAGCTGCTGGAGGCTCCTTTGATTCCTAGACCTGG	2100
2041	Db	TGCTTGCTCAGCTCAGGCCCTCCTAGCTGCTGGAGGCTCCTTTGATTCCTAGACCTGG	2100
2101	QY	AAAAGTGTCCTTAGCAGAGCCCTGSCAGGGGCTCAGAGCTGGGATTTCTTGCCCTGGA	2160
2101	Db	AAAAGTGTCCTTAGCAGAGCCCTGSCAGGGGCTCAGAGCTGGGATTTCTTGCCCTGGA	2160
2161	QY	ACAAGGACCTCGAGAAATGTTTTTGCTGGCATGATGTGTGCTCAGGAGCCCTCTGGGC	2220
2161	Db	ACAAGGACCTCGAGAAATGTTTTTGCTGGGATGATGTGTGCTCAGGAGCCCTCTGGGC	2220
2221	QY	ATCGTTCCTCTGCCTTTGGTAGTCCAGGACAGGCCAATGATGTTCTCAGTAGCCT	2280
2221	Db	ATCGTTCCTCTGCCTTTGGTAGTCCAGGACAGGCCAATGATGTTCTCAGTAGCCT	2280
2281	QY	TATCATTCAGAGTGCTCTCTAGCTGCACAAATGATTCACAAGAGATCACCCCAAGGA	2340
2281	Db	TATCATTCAGAGTGCTCTCTAGCTGCACAAATGATTCACAAGAGATCACCCCAAGGA	2340
2341	QY	TTATTTCTGAAGGTGTTTTTTCTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTTTT	2400
2341	Db	TTATTTCTGAAGGTGTTTTTTCTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTTTT	2400
2401	QY	TTTTTTTTGCATGACAGTGTGTATTTAGGACCTTCCAAGGAAAGGATGCTGTACC	2460
2401	Db	TTTTTTTTGCATGACAGTGTGTATTTAGGACCTTCCAAGGAAAGGATGCTGTACC	2460
2461	QY	AGTGGTCCCTGGGTGCTGGCCTCCAGTGTCCACCTCTTTCCACCCACCTTGGCTCC	2520
2461	Db	AGTGGTCCCTGGGTGCTGGCCTCCAGTGTCCACCTCTTTCCACCCACCTTGGCTCC	2520
2521	QY	TTTGCCATCTTGATGCTGAGGTTTCCTGTTTGTGTGAGATCAGGTTGTTGTGTAAGA	2580
2521	Db	TTTGCCATCTTGATGCTGAGGTTTCCTGTTTGTGTGAGATCAGGTTGTTGTGTAAGA	2580
2581	QY	AAGGAAAGGGCTCTCTGATGGCTTTGCCACAAGCTTACTCTGGGTTTCAGTCTGACAGG	2640
2581	Db	AAGGAAAGGGCTCTCTGATGGCTTTGCCACAAGCTTACTCTGGGTTTCAGTCTGACAGG	2640
2641	QY	CCACCACAGTTCCCATCAGCACTGHTCTCCATGACAGAGTTGCTGGGTCCCATGTCCAGC	2700
2641	Db	CCACCACAGTTCCCATCAGCACTGHTCTCCATGACAGAGTTGCTGGGTCCCATGTCCAGC	2700

QY	2701	TG	CCTCTTTGGCTTCATGGGTTTTTTCGTCTCTGCCCCACACCCACATGTGCAATCCT	2760
DB	2701	TG	CCTCTTTGGCTTCATGGGTTTTTTCGTCTCTGCCCCACACCCACATGTGCAATCCT	2760
QY	2761	CA	GATTTGCTCGATTCATTTCTCGTGACCTCCCTGCTGTCCCTTGGGATTCCTACATT	2820
DB	2761	CA	GATTTGCTCGATTCATTTCTCGTGACCTCCCTGCTGTCCCTTGGGATTCCTACATT	2820
QY	2821	CTT	CCTGTGTGGGCCCATAGCTGTTGTCTAACAGGTAAGAAATGAAATCTATTGA	2880
DB	2821	CTT	CCTGTGTGGGCCCATAGCTGTTGTCTAACAGGTAAGAAATGAAATCTATTGA	2880
QY	2881	CTGG	CCCCAGAAATCATAAATGGCTGCAGACAGTTGTTTCTGTGCTCTGTCTACCC	2940
DB	2881	CTGG	CCCCAGAAATCATAAATGGCTGCAGACAGTTGTTTCTGTGCTCTGTCTACCC	2940
QY	2941	CCACT	CCAGTACATAAATCTATCTATGTACTGTGTAGAGCCATTCTATATGCTGAATGTTCTG	3000
DB	2941	CCA	CTCCAGTACATAAATCTATCTATGTACTGTGTAGAGCCATTCTATATGCTGAATGTTCTG	3000
QY	3001	CTGTT	GCAACTTCGCCAGGTATTAGCCAGTGTGTCGCAACAGATTTTCGGGGACAAC	3060
DB	3001	CTGTT	GCAAACTTCGCCAGGTATTAGCCAGTGTGTCGCAACAGATTTTCGGGGACAAC	3060
QY	3061	AGAAT	GACTCAGACCAAGATGGATAGGATGGTTAGGGCTTTTGTTCTGTGCTGTTTTTCTT	3120
DB	3061	AGAAT	GACTCAGACCAAGATGGATAGGATGGTTAGGGCTTTTGTTCTGTGCTGTTTTTCTT	3120
QY	3121	TGA	CTAGTCAATTGTCCTGAGGTCCCTTCATCTTCCATACCTPAGGCCACTCTTTTAGCC	3180
DB	3121	TGA	CTAGTCAATTGTCCTGAGGTCCCTTCATCTTCCATACCTPAGGCCACTCTTTTAGCC	3180
QY	3181	CTT	ACCTTAAATCTCTCAGATAGTGGTTTCACAAGAATGTTTAAGTACTGAATCATGTG	3240
DB	3181	CTT	ACCTTAAATCTCTCAGATAGTGGTTTCACAAGAATGTTTAAGTACTGAATCATGTG	3240
QY	3241	TG	CTGACACAGAGATGGCAAAATGAAATGGCACACCAATTTCTCCTTCTCTGCCCCAGGG	3300
DB	3241	TG	CTGACACAGAGATGGCAAAATGAAATGGCACACCAATTTCTCCTTCTCTGCCCCAGGG	3300
QY	3301	CAG	TACATGATCTGCATCAGATTTGCCCTGCTATTTCTCTGGTGATCTCTTACATCTA	3360
DB	3301	CAG	TACCATCTGATCTGCATCAGATTTGCCCTGCTATTTCTCTGGTGATCTCTTACATCTA	3360
QY	3361	GGT	GCCTCAAGCAGCTGTGTGAGTCTCGCATCTCGCCATCTCTGGCTGAATACTGC	3420
DB	3361	GGT	GCCTCAAGCAGCTGTGTGAGTCTCGCATCTCGCCATCTCTGGCTGAATACTGC	3420
QY	3421	TGT	CCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTGTGAATCCCTCTTGTCTGGAG	3480
DB	3421	TGT	CCTGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTGTGAATCCCTCTTGTCTGGAG	3480
QY	3481	TAG	TCCCTGCCCTTCTTCTGCTCCAGTAGGCCCTTTTCTTACCCACGCCCTTGTGCCAGG	3540
DB	3481	TAG	TCCCTGCCCTTCTTCTGCTCCAGTAGGCCCTTTTCTTACCCACGCCCTTGTGCCAGG	3540
QY	3541	CTA	AGCTGGTACAAGAGCTGCCAATCACAGAGTTTGTAGGCGAGAGGTCAGGGA	3600
DB	3541	CTA	AGCTGGTACAAGAGCTGCCAATCACAGAGTTTGTAGGCGAGAGGTCAGGGA	3600
QY	3601	AG	GCAGAGTATGCACTTCCCTTGAAGAGAGGGGAAGCCCTACAGTGGCCACA	3660
DB	3601	AG	GCAGAGTATGCACTTCCCTTGAAGAGAGGGGAAGCCCTACAGTGGCCACA	3660
QY	3661	TAA	TGCTGACTCACACTTCACTCTTAATGCTCTGGAGGACCTGGAGCTGCTG	3720
DB	3661	TAA	TGCTGACTCACACTTCACTCTTAATGCTCTGGAGGACCTGGAGCTGCTG	3720
QY	3721	GAT	CCAGTGTGGTGTAGGAGGCCACAGTGAGAGGTTGGCCCCAGCTGGTGTTCCTCA	3780
DB	3721	GAT	CCAGTGTGGTGTAGGAGGCCACAGTGAGAGGTTGGCCCCAGCTGGTGTTCCTCA	3780

QY 3781 GGTGAGAAATGTGGGCCCCCAGGAGGTGCTGCTCAGAGTCCATCCATGCTCA 3840
DB 3781 GGTGAGAAATGTGGGCCCCCAGGAGGTGCTGCTCAGAGTCCATCCATGCTCA 3840
QY 3841 GACCTTCAGGCCAGTCTGAGATGAGGTCCCTPACCTTTTCTCTCTTCATTGACCAA 3900
DB 3841 GACCTTCAGGCCAGTCTGAGATGAGGTCCCTPACCTTTTCTCTCTTCATTGACCAA 3900
QY 3901 TCAACCAATCACTACAGTCTGCTGCTTCTGCTTCCAAAGTAGCCAGGTCTGGGCA 3960
DB 3901 TCAACCAATCACTACAGTCTGCTGCTTCTGCTTCCAAAGTAGCCAGGTCTGGGCA 3960
QY 3961 GATGCGGGGAGGTGCTATCCATGATGAGTGAAGCCAGTGTCTTCCCTCACTGGGTGTC 4020
DB 3961 GATGCGGGGAGGTGCTATCCATGATGAGTGAAGCCAGTGTCTTCCCTCACTGGGTGTC 4020
QY 4021 CACACTGTGACCTCAGTTTATAGGACCCCAAGATCTGTGTGTTTCTTAGATTGCTAGC 4080
DB 4021 CACACTGTGACCTCAGTTTATAGGACCCCAAGATCTGTGTGTTTCTTAGATTGCTAGC 4080
QY 4081 TTTTCTCCAGGGAGCCACAGAGGTGAAGCTCAAGAGCGCATGCTCTGCTAATAGTAA 4140
DB 4081 TTTTCTCCAGGGAGCCACAGAGGTGAAGCTCAAGAGCGCATGCTCTGCTAATAGTAA 4140
QY 4141 ATTGTTTTCAGGGCCTTGCTGAGGAGTCTGATTTCTGTAAGATGTTATTTAACTCTCC 4200
DB 4141 ATTGTTTTCAGGGCCTTGCTGAGGAGTCTGATTTCTGTAAGATGTTATTTAACTCTCC 4200
QY 4201 AGAGCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4260
DB 4201 AGAGCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4260
QY 4261 GAGCTCAGGACATTTCTGAGGAGTCTGATTTCTGTAAGATGTTATTTAACTCTCC 4320
DB 4261 GAGCTCAGGACATTTCTGAGGAGTCTGATTTCTGTAAGATGTTATTTAACTCTCC 4320
QY 4321 TCAACCCCAATCAAGTGGGCTGAGGCTGACCCGAGGAGCCAGTGGAGTGGCTGGTGTGTC 4380
DB 4321 TCAACCCCAATCAAGTGGGCTGAGGCTGACCCGAGGAGCCAGTGGAGTGGCTGGTGTGTC 4380
QY 4381 CAGGGGGAGGAGGAGGCTGAGGCTGAGTATTTCTGTAAGATGTTATTTAACTCTCC 4440
DB 4381 CAGGGGGAGGAGGAGGCTGAGGCTGAGTATTTCTGTAAGATGTTATTTAACTCTCC 4440
QY 4441 CTTGCAAGCAGACAGAGTGGTCAACCCAGGAGCAGCCAGGACCTGCTCTCTGCTCTT 4500
DB 4441 CTTGCAAGCAGACAGAGTGGTCAACCCAGGAGCAGCCAGGACCTGCTCTCTGCTCTT 4500
QY 4501 CTTGGGGAGGAGGAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
DB 4501 CTTGGGGAGGAGGAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
QY 4561 CACTCAGACTTGTGTAAGTGCACCTGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4620
DB 4561 CACTCAGACTTGTGTAAGTGCACCTGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4620
QY 4621 AGCAGGAGTGAATTTGATGATTTTAAACCAATTAAGAGTCTGCTGCTGCTGCTGCT 4680
DB 4621 AGCAGGAGTGAATTTGATGATTTTAAACCAATTAAGAGTCTGCTGCTGCTGCTGCT 4680
QY 4681 AAAAAAAAAAAAAAAAAA 4697
DB 4681 AAAAAAAAAAAAAAAAAA 4697

RESULT 6

US-10-616-187-17
; Sequence 17, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.

APPLICANT: Ariona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/616.187
CURRENT FILING DATE: 2003-07-09
PRIOR APPLICATION NUMBER: US/09/616.289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 4697
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(1592)
US-10-616-187-17

Query Match 100.0%; Score 4697; DB 16; Length 4697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAAAGAGCAGCCCAAGCAACCGGAGCAGGACCCGAGGAGCCAGGAGCCGAGCC 60
DB 1 CAAAAGAGCAGCCCAAGCAACCGGAGCAGGACCCGAGGAGCCAGGAGCCGAGCC 60
QY 61 AGCGGGCTCTGAGTGAAGAGAGAGTCCCGGAGCAGCAGGCTCTCTGGAAGCCGG 120
DB 61 AGCGGGCTCTGAGTGAAGAGAGAGTCCCGGAGCAGCAGGCTCTCTGGAAGCCGG 120
QY 121 AGGGGGCTCAAGCAGACAGCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGA 180
DB 121 AGGGGGCTCAAGCAGACAGCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGA 180
QY 181 GCCGCCAATCTGGAAGACATCTGAGACATCTGTGTGGAATATACAGAGGGGGCCCG 240
DB 181 GCCGCCAATCTGGAAGACATCTGAGACATCTGTGTGGAATATACAGAGGGGGCCCG 240
QY 241 GCGAGGATGGGGCAGCGGTGAGCCGCTGAACCCGAGAGATGAGAGAGTCCCGGACCT 300
DB 241 GCGAGGATGGGGCAGCGGTGAGCCGCTGAACCCGAGAGATGAGAGAGTCCCGGACCT 300
QY 301 ATGTGGCAAGAAATGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGAGACCT 360
DB 301 ATGTGGCAAGAAATGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGAGACCT 360
QY 361 CCAAGGGGATCCAAACACAGAGAGATCCGGAGAGTCCGAGAGTCCGAGAGCCGAGAC 420
DB 361 CCAAGGGGATCCAAACACAGAGAGATCCGGAGAGTCCGAGAGTCCGAGAGCCGAGAC 420
QY 421 ATCGAAGGCCACAGGAG 480
DB 421 ATCGAAGGCCACAGGAG 480
QY 481 TGCAGACATTGAATCTCTGAGTACCCAGAGAGAGTGGTGTCTGTGCAAGAGT 540
DB 481 TGCAGACATTGAATCTCTGAGTACCCAGAGAGAGTGGTGTCTGTGCAAGAGT 540
QY 541 ATGCTGAATCTGCTGGAGAGCAGCCGAAATTCACAGAGAGAGATGAAGCTCTCTACAGAAA 600
DB 541 ATGCTGAATCTGCTGGAGAGCAGCCGAAATTCACAGAGAGAGATGAAGCTCTCTACAGAAA 600
QY 601 AGCAGAGCCAGCTGGTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

601	Db	AGCAGAGCCAGCTGGTGTCAAGAGAAGGACCAACTGGCGCGGTGAGCAGACAGCAAGAGCGCGTCC	660
661	QY	TGSCCGCAGCAGCTTGAGAGCGCTATGCGGTGAGCTGCAGCGGCAACAACCGTCCCTCA	720
661	Db	TGSCCGCAGCAGCTTGAGAGCGCTATGCGGTGAGCTGCAGCGGCAACAACCGTCCCTCA	720
721	QY	AGGAAGAAGGTGTGACGCGCGCCCGGAGGAGAGGAGCGCAAGAGGAGTGACTCGC	780
721	Db	AGGAAGAAGGTGTGACGCGCGCCCGGAGGAGAGGAGCGCAAGAGGAGTGACTCGC	780
781	QY	ACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGACACAATGAGCGCACT	840
781	Db	ACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGACACAATGAGCGCACT	840
841	QY	CCAAGCTGGCGCCAAGAGAACATGAGAGCTGGCTCAGAGGCTCAAGAAAGCTGATTGAGCAGT	900
841	Db	CCAAGCTGGCGCCAAGAGAACATGAGAGCTGGCTCAGAGGCTCAAGAAAGCTGATTGAGCAGT	900
901	QY	ATGAGCTGGCGGAGGAGCATATCGAACAAGTCTTCAACACAAAGACCTTACACAGCAGC	960
901	Db	ATGAGCTGGCGGAGGAGCATATCGAACAAGTCTTCAACACAAAGACCTTACACAGCAGC	960
961	QY	TGCTGGATGCCAAGCTCCAGCAGGCCAGAGAGTGTAAAGAGGAGGAGAGCGGCACC	1020
961	Db	TGCTGGATGCCAAGCTCCAGCAGGCCAGAGAGTGTAAAGAGGAGGAGAGCGGCACC	1020
1021	QY	AGCGGAGAAGGATTTTCTCCTGAAAGAGCGATGAGTCCCAGAGGATGTGTGAGCTGA	1080
1021	Db	AGCGGAGAAGGATTTTCTCCTGAAAGAGCGATGAGTCCCAGAGGATGTGTGAGCTGA	1080
1081	QY	TGAAGCAGCAAGAGAGCCACCTCAAGCAACAGCTTGCCCTTATACAGAGAAGTTTGAGG	1140
1081	Db	TGAAGCAGCAAGAGAGCCACCTGAGCAACAGCTTGCCCTTATACAGAGAAGTTTGAGG	1140
1141	QY	AGTTCCAGAACACACTTTTCCAAAAGCAGCGAGTATTCACCACTTCAAGACGAGATGG	1200
1141	Db	AGTTCCAGAACACACTTTTCCAAAAGCAGCGAGTATTCACCACTTCAAGACGAGATGG	1200
1201	QY	AAAAGTGACTAGAGATCAGAAGCTGAGAGAAAGAAACACCATGTACCGGTCCCGGT	1260
1201	Db	AAAAGTGACTAGAGATCAGAAGCTGAGAGAAAGAAACCACTGTACCGGTCCCGGT	1260
1261	QY	GGAGAGCAGCAACAAGGCCCTGCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAG	1320
1261	Db	GGAGAGCAGCAACAAGGCCCTGCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAG	1320
1321	QY	AACGTGAGGCCCTGCAGTAAATCAACGGCTGGAGAGCTGTCCGGGCACTCAGA	1380
1321	Db	AACGTGAGGCCCTGCAGTAAATCAACGGCTGGAGAGCTGTCCGGGCACTCAGA	1380
1381	QY	CAGAGCCCAATGACTTGAACAAGAGGATACAGACCTGAGTGTGGTGGGCACAGGCTCCC	1440
1381	Db	CAGAGCCCAATGACTTGAACAAGAGGATACAGACCTGAGTGTGGTGGGCACAGGCTCCC	1440
1441	QY	TCAGTGAAGTGGCCCTGAGAGAGGCAGAGGGCCCTGGGGCTCAAGCAACCAGTCCC	1500
1441	Db	TCAGTGAAGTGGCCCTGAGAGAGGCAGAGGGCCCTGGGGCTCAAGCAACCAGTCCC	1500
1501	QY	CCAGGCTCACAAGACCGCCTTGCTTACCACAGAGCACCGAGCACAGAGCAATCAGGCCAGA	1560
1501	Db	CCAGGCTCACAAGACCGCCTTGCTTACCAGGAGCACCGAGCACAGAGCAATCAGGCCAGA	1560
1561	QY	CTGGGCCCTCAAGAGCCCACTTCGCCCAGGGCCTTAGAGAGCCTGGTGTGGGTATGCTGG	1620
1561	Db	CTGGGCCCTCAAGAGCCCACTTCGCCCAGGGCCTTAGAGAGCCTGGTGTGGGTATGCTGG	1620
1621	QY	GAAGGAGCGCAGCCAGCCAGCGCTTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCCA	1680
1621	Db	GAAGGAGCGCAGCCAGCCAGCGCTTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCCA	1680
1681	QY	TTGCTGAAGCCAGAGTTCCTGACCTGGCTGGCACTTGCGCACTTGCAATTTTGGATTTT	1740

Db	1681	TTGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTCGCACTTCCAAATTTTGGATTTT	1741
Qy	1741	GTGGGTCAAGTTTACGTACATAGGGCAATTTGCAAGGCCTTGCAATGCAATTTATACCTG	1801
Db	1741	GTGGGTCAAGTTTACGTACATAGGGCAATTTGCAAGGCCTTGCAATGCAATTTATACCTG	1801
Qy	1801	TAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAAGTGCCTTG	1861
Db	1801	TAAGTGTACAGTGGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAAGTGCCTTG	1861
Qy	1861	TCGTGTAGCTGAAGAGTCTTTGAGAGGGCTGTGCATCTGTAGCTGCCATCACAGTGAAGTTG	1921
Db	1861	TCGTGTAGCTGAAGAGTCTTTGAGAGGGCTGTGCATCTGTAGCTGCCATCACAGTGAAGTTG	1921
Qy	1921	GCAGAAGTGACTTGAGCAATTTCTCTGTCTGAATTTGAGGCTCAGACCCCTCCCTGCCCTTT	1981
Db	1921	GCAGAAGTGACTTGAGCAATTTCTCTGTCTGAATTTGAGGCTCAGACCCCTCCCTGCCCTTT	1981
Qy	1981	CAGAGCTCAAAAAGAAATACACCAAGGTCTTGACTGCAATTTGTCTGTGTGAGCAGGGCT	2041
Db	1981	CAGAGCTCAAAAAGAAATACACCAAGGTCTTGACTGCAATTTGTCTGTGTGAGCAGGGCT	2041
Qy	2041	TGCTTGTGCTCAGCTCAGGCCCTCCTAGCTGCTTGAGAGGCTCCTTTGATCTCTAGACCTGG	2101
Db	2041	TGCTTGTGCTCAGCTCAGGCCCTCCTAGCTGCTTGAGAGGCTCCTTTGATCTCTAGACCTGG	2101
Qy	2101	AAAAGGTGTCCTAGGCAGAGCCCTGCAGGGGCTCAGAGCTGGGATTTCCCTGCCCTGGA	2161
Db	2101	AAAAGGTGTCCTAGGCAGAGCCCTGCAGGGGCTCAGAGCTGGGATTTCCCTGCCCTGGA	2161
Qy	2161	ACAAAGGACCTGGAGAAATTTTTTGCTGGGATGATGTGTGTGCTAGGAGCCCTTGCGC	2221
Db	2161	ACAAAGGACCTGGAGAAATTTTTTGCTGGGATGATGTGTGTGCTAGGAGCCCTTGCGC	2221
Qy	2221	ATCCCTTCCCTGCCCTTGCTAGTGCAGGACCCAGGCCAATGATGCTCTCAGTAGCCCT	2281
Db	2221	ATCCCTTCCCTGCCCTTGCTAGTGCAGGACCCAGGCCAATGATGCTCTCAGTAGCCCT	2281
Qy	2281	TATCATTTACAGGTGCTCTCTAGCTTGCAAAATGATGCAAGAGATCACCCAAAGGA	2341
Db	2281	TATCATTTACAGGTGCTCTCTAGCTTGCAAAATGATGCAAGAGATCACCCAAAGGA	2341
Qy	2341	TTATTTCTGAAGGTGTTTTTTCTTTATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	2401
Db	2341	TTATTTCTGAAGGTGTTTTTTCTTTATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	2401
Qy	2401	TTTTTTTGCATACACAGTGTGTTGTAATGAGGACCTTCCAAGGAAAAGGATGCTGPACC	2461
Db	2401	TTTTTTTGCATACACAGTGTGTTGTAATGAGGACCTTCCAAGGAAAAGGATGCTGPACC	2461
Qy	2461	AGTGTGCTGGGTGCTTGGCCCTCAGTGTCCACCTCCTTCCACCCCCACATTTGGCTCC	2521
Db	2461	AGTGTGCTGGGTGCTTGGCCCTCAGTGTCCACCTCCTTCCACCCCCACATTTGGCTCC	2521
Qy	2521	TTTGGCAATCTTGATGCTCAGGTTTTCCCTGTGTTGGTGAGATCAGGTGTGTTGTGTAAGA	2581
Db	2521	TTTGGCAATCTTGATGCTCAGGTTTTCCCTGTGTTGGTGAGATCAGGTGTGTTGTGTAAGA	2581
Qy	2581	AAGGAAAGGGCTTCTGATGGCTTTGCCACAGCTTACCTGTGGGTTTTCAGTCTCTGAGG	2641
Db	2581	AAGGAAAGGGCTTCTGATGGCTTTGCCACAGCTTACCTGTGGGTTTTCAGTCTCTGAGG	2641
Qy	2641	CCACCACAGTTCCCATCAGCACTGTCTCCATGCAGCAGTTGCTGGGTCCTATGTCAGC	2701
Db	2641	CCACCACAGTTCCCATCAGCACTGTCTCCATGCAGCAGTTGCTGGGTCCTATGTCAGC	2701
Qy	2701	TGCTCTTTGGCTTCATGGGTTTTCTGCTTCTGCTCCGCCCCACCCCAATGTCATCTCT	2761
Db	2701	TGCTCTTTGGCTTCATGGGTTTTCTGCTTCTGCTCCGCCCCACCCCAATGTCATCTCT	2761
Qy	2761	CAAGATTTGCTGATTTCTATTTCTGGCACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG	2821
Db	2761	CAAGATTTGCTGATTTCTATTTCTGGCACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG	2821

2821 QY CTTCTGTGGGGCCCATAGCTGTTGTTCTAAACAGGTAAAGAAATGAATTAATGA 2880
2821 Db CTTCTGTGTGGGGCCCATAGCTGTTGTTCTAAACAGGTAAAGAAATGAATTAATGA 2880
2881 QY CTGGGCCCCAGAAATCATAAATAGCTGTCAGACAGTTGTTCTGTGCTGCTTACCC 2940
2881 Db CTGGGCCCCAGAAATCATAAATAGCTGTCAGACAGTTGTTCTGTGCTGCTTACCC 2940
2941 QY CCACCTCAGTACATACTATGTAAGTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
2941 Db CCACCTCAGTACATACTATGTAAGTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
3001 QY CTGTTGCAAACTTGGCCAGGTATAGCCAGTGTGTCAGACAGTTTTCGGGGACAAC 3060
3001 Db CTGTTGCAAACTTGGCCAGGTATAGCCAGTGTGTCAGACAGTTTTCGGGGACAAC 3060
3061 QY AGAATGACTCAGACCAAGATAGGATAGGTTAGGGCTTTGCTTCTGCTGTTTTCTT 3120
3061 Db AGAATGACTCAGACCAAGATAGGATAGGTTAGGGCTTTGCTTCTGCTGTTTTCTT 3120
3121 QY TGAACCTAGTCAATGTCCTGAGTCCCTTCATCTTCCATACCTAGCCCACTCTTTAGCC 3180
3121 Db TGAACCTAGTCAATGTCCTGAGTCCCTTCATCTTCCATACCTAGCCCACTCTTTAGCC 3180
3181 QY CTTACCTTAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
3181 Db CTTACCTTAATCTCTCAGATAAGTTGGTTTCAAAAGAAATGTTAAGTACTGAATCATGTG 3240
3241 QY TGAACCTAGACAGATAGGCAATGAATGAGCAACCAATTTCTCTCTGCTGCCAGGG 3300
3241 Db TGAACCTAGACAGATAGGCAATGAATGAGCAACCAATTTCTCTCTGCTGCCAGGG 3300
3301 QY CAGTACCACTGATCTGCATCAGATGTCCTGATTTCTCTGCTGATCTCTTCAATCTA 3360
3301 Db CAGTACCACTGATCTGCATCAGATGTCCTGATTTCTCTGCTGATCTCTTCAATCTA 3360
3361 QY GGTGCCCTCAAGCAGTGTGTAGTGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGC 3420
3361 Db GGTGCCCTCAAGCAGTGTGTAGTGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGC 3420
3421 QY TGTCTGTGAAGTGTTCCTCATGACCTTTTCTCTCCCTTTGAAATCCCTTTGCTGGAG 3480
3421 Db TGTCTGTGAAGTGTTCCTCATGACCTTTTCTCTCCCTTTGAAATCCCTTTGCTGGAG 3480
3481 QY TAGTCCCTGCTTCTCTCTCCAGTACCTTTTCTCTCCCTTTGAAATCCCTTTGCTGGAG 3540
3481 Db TAGTCCCTGCTTCTCTCTCCAGTACCTTTTCTCTCCCTTTGAAATCCCTTTGCTGGAG 3540
3541 QY CTAAGCTGTGTAGAGTGTGCAAGTGTGTCAGAGTGTGTCAGAGAGAGAGTGTGAGGA 3600
3541 Db CTAAGCTGTGTAGAGTGTGCAAGTGTGTCAGAGTGTGTCAGAGAGAGAGTGTGAGGA 3600
3601 QY AGAGGAGAGAGTATGACCTTCCCTTCAAGAGAGGGAAGGCTTACAGTGGGCCACA 3660
3601 Db AGAGGAGAGAGTATGACCTTCCCTTCAAGAGAGGGAAGGCTTACAGTGGGCCACA 3660
3661 QY TAATTCCTGACTCAGTACCTTCAAGTACCTTTAATGCTGTGAGGAGTGTGAGTGTG 3720
3661 Db TAATTCCTGACTCAGTACCTTCAAGTACCTTTAATGCTGTGAGGAGTGTGAGTGTG 3720
3721 QY GATCCAGTGTGTGTGTAGAGGCAAGTGTGAGAGTGTGTCAGAGTGTGCTGCTTCCCA 3780
3721 Db GATCCAGTGTGTGTGTAGAGGCAAGTGTGAGAGTGTGTCAGAGTGTGCTGCTTCCCA 3780
3781 QY GGTGAGAAATGTGGGCCAGGCAAGTGTGCTTGTCTCAGCTCCATCATGCTCTA 3840
3781 Db GGTGAGAAATGTGGGCCAGGCAAGTGTGCTTGTCTCAGCTCCATCATGCTCTA 3840
3841 QY GACCTTCAGGCCAGTCTGAGATGAGGTTCCCTTACCTTTCTCTCTCATTTGACCAAA 3900
3841 Db GACCTTCAGGCCAGTCTGAGATGAGGTTCCCTTACCTTTCTCTCTCATTTGACCAAA 3900

3901 QY TCAACCAATCACTACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
3901 Db TCAACCAATCACTACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
3961 QY GATCAGGGAGGTGCTTATCCATGAGTGAAGGCAAGTGTCTCTCTCTCTCTCTCTCTCT 4020
3961 Db GATCAGGGAGGTGCTTATCCATGAGTGAAGGCAAGTGTCTCTCTCTCTCTCTCTCTCT 4020
4021 QY CACACTTGTGACCTCTCAGTTTATAGGACCAAGATCTCTGTGTGTGTCTCTAGATTGCTAGC 4080
4021 Db CACACTTGTGACCTCTCAGTTTATAGGACCAAGATCTCTGTGTGTGTCTCTAGATTGCTAGC 4080
4081 QY TTTTCTCCAGGGGACACACAGCAGTGAAGTCAAGAGCCGATGGCTCTCTCTAATAGTAA 4140
4081 Db TTTTCTCCAGGGGACACACAGCAGTGAAGTCAAGAGCCGATGGCTCTCTCTAATAGTAA 4140
4141 QY ATTGTTTTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACAGATTCTGAGAGGTGTC 4200
4141 Db ATTGTTTTTTCAGGGCTTGTCCAGCTGAGAGCTTCATGTCCACAGATTCTGAGAGGTGTC 4200
4201 QY AGCAGCACTTTTTTTTTTTTATTTTATTTTGTGTTTTCATGAGTTATCGAACCATGGCT 4260
4201 Db AGCAGCACTTTTTTTTTTTTATTTTATTTTGTGTTTTCATGAGTTATCGAACCATGGCT 4260
4261 QY GAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTCC 4320
4261 Db GAGCTCAGGCACTTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTATTTAAACCTCC 4320
4321 QY TCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGAGGCGCAGTGGAGCTGCTTGTGTC 4380
4321 Db TCCACCCCATCACGGTGGCCCTGAGGGCTGACCCGAGGCGCAGTGGAGCTGCTTGTGTC 4380
4381 QY CACGGGAGGCGCCAGAGCTGCTGAGCTGATTTCTCAGCTGTGCTGCCAGGCTTTCCGC 4440
4381 Db CACGGGAGGCGCCAGAGCTGCTGAGCTGATTTCTCAGCTGTGCTGCCAGGCTTTCCGC 4440
4441 QY CTTGCAAGCAGACAGAGTGGTACCCAGGAGCAGCCAGGACCTGCTCTCTTGTGCCCTT 4500
4441 Db CTTGCAAGCAGACAGAGTGGTACCCAGGAGCAGCCAGGACCTGCTCTCTTGTGCCCTT 4500
4501 QY CCTGGGAAAGAGAGTGGCTTCTGTCCCTGTAACTGCTTCTTATGCGCCCAACCGGC 4560
4501 Db CCTGGGAAAGAGAGTGGCTTCTGTCCCTGTAACTGCTTCTTATGCGCCCAACCGGC 4560
4561 QY CACTCAGACTTGTGTGAAGCTGCACTGCGAGCTTTTGTCTCTCTTGGGTATTCAAC 4620
4561 Db CACTCAGACTTGTGTGAAGCTGCACTGCGAGCTTTTGTCTCTCTTGGGTATTCAAC 4620
4621 QY AGCCAGGAGCTTGAATTTGATTTTAAACCAATTAATAAAGAGTCTGTGCTTA 4680
4621 Db AGCCAGGAGCTTGAATTTGATTTTAAACCAATTAATAAAGAGTCTGTGCTTA 4680
4681 QY AAAAAAAAAAAAAAAAAA 4697
4681 Db AAAAAAAAAAAAAAAAAA 4697

RESULT 7

US-10-276-774-784

; Sequence 784, Application US/10276774

; Publication No. US20040053245A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-030

; CURRENT APPLICATION NUMBER: US/10/276,774

; PRIOR FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: 09/560,875

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 09/496,914

; PRIOR FILING DATE: 2000-02-03

; NUMBER OF SEQ ID NOS: 2700

SOFTWARE: Custom
SEQ ID NO 784
LENGTH: 2523
TYPE: DNA
ORGANISM: Homo sapiens
US-10-276-774-784

Query Match 33.8%; Score 1589; DB 13; Length 2523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	CAAAAAGCAGCCAGGACAAACCGAAGCAGGAGCCAGGAGCCAGGAGCCAGCC	60
Db	165	CAAAAAGCAGCCAGGACAAACCGAAGCAGGAGCCAGGAGCCAGGAGCCAGCC	224
QY	61	AGCGCGCTCTCAGTACGAGCAGAGGTCGCGGAGCAGCAGGCTCTCGAAGCCGG	120
Db	225	AGCGCGCTCTCAGTACGAGCAGAGGTCGCGGAGCAGCAGGCTCTCGAAGCCGG	284
QY	121	AGCGCGCTCAAGCCAGAGCGCTCAGTCTGGGGCCCTTGGTGATCTCTGAGGAGCTGA	180
Db	285	AGCGCGCTCAAGCCAGAGCGCTCAGTCTGGGGCCCTTGGTGATCTCTGAGGAGCTGA	344
QY	181	GGCGCGAAGTGAAGCAGCAGTACTGAGCAGTACTGTGAGCAGTAAACCGAGGAGCCCGG	240
Db	345	GGCGCGAAGTGAAGCAGCAGTACTGAGCAGTACTGTGAGCAGTAAACCGAGGAGCCCGG	404
QY	241	GGAGGATGGGCGACAGGCTGAGCGGCTGAACCCGAGATCGAGAGAGTCCCGGACCT	300
Db	405	GGAGGATGGGCGACAGGCTGAGCGGCTGAACCCGAGATCGAGAGAGTCCCGGACCT	464
QY	301	ATGTGCAAGGAAATGGGAGCTGAAACCACTCCAGTAGTCTATGGAGAGAGGACCT	360
Db	465	ATGTGCAAGGAAATGGGAGCTGAAACCACTCCAGTAGTCTATGGAGAGAGGACCT	524
QY	361	CAAGGGGATCCAAACAGAGAGATCCGCGAGATCGAGGTCGGAGACCCGAGACC	420
Db	525	CAAGGGGATCCAAACAGAGAGATCCGCGAGATCGAGGTCGGAGACCCGAGACC	584
QY	421	ATCGAAGGCGACAGGAGAGAAAGAAAGCAAGGTTTGGGAGAGAGATCACTTGTCTGA	480
Db	585	ATCGAAGGCGACAGGAGAGAAAGAAAGCAAGGTTTGGGAGAGAGATCACTTGTCTGA	644
QY	481	TGCAGACATGATCTCTGAGTACCCAGAGAGAGAGTCTGCTCTGTCAGAGAGT	540
Db	645	TGCAGACATGATCTCTGAGTACCCAGAGAGAGAGTCTGCTCTGTCAGAGAGT	704
QY	541	ATGCTGAACTGTGAGGAGCAGCCGAAATTCAGAGAGCAGATGAAGCTCTTACAGAAA	600
Db	705	ATGCTGAACTGTGAGGAGCAGCCGAAATTCAGAGAGCAGATGAAGCTCTTACAGAAA	764
QY	601	AGCAGGCGAGCTGGTGCAAGAGAGAGCAGCAGTGGCGGTGAGCAGCAGAGGCGCTCC	660
Db	765	AGCAGGCGAGCTGGTGCAAGAGAGAGCAGCAGTGGCGGTGAGCAGCAGAGGCGCTCC	824
QY	661	TGGCCCGCAGCAAGCTTGAGAGCTATGCCGTGAGCTGCGGCGCAACCGCTCCCTCA	720
Db	825	TGGCCCGCAGCAAGCTTGAGAGCTATGCCGTGAGCTGCGGCGCAACCGCTCCCTCA	884
QY	721	AGGAAGAGGTGTGACGCGGCGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTCCTC	780
Db	885	AGGAAGAGGTGTGACGCGGCGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTCCTC	944
QY	781	ACTTCCAGGTGACACTGAAATGATTCAGCTGAGATGGAACAGCAGCAATGAGCGCACT	840
Db	945	ACTTCCAGGTGACACTGAAATGATTCAGCTGAGATGGAACAGCAGCAATGAGCGCACT	1004
QY	841	CCAAGCTGCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTGAGCAGT	900
Db	1005	CCAAGCTGCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTGAGCAGT	1064
QY	901	ATGAGCTGCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAA	960

Db	1065	ATGAGCTGGCGGAGGAGCATATCGAATAAGTCTTCAAAACAGAGGACCTTACACAGCAGC	1124
QY	961	TGCTGGATGCCAAGCTCCAGCAGGAGGAGATGCTAAAGGAGGAGAGAGAGAGCAGCACC	1020
Db	1125	TGCTGGATGCCAAGCTCCAGCAGGAGGAGATGCTAAAGGAGGAGAGAGAGAGCAGCACC	1184
QY	1021	AGCGGAGAGAGAGATTTTCTCTGAAAGAGCAGTACAGTCCAGAGAGATGTGTGAGCTGA	1080
Db	1185	AGCGGAGAGAGAGATTTTCTCTGAAAGAGCAGTACAGTCCAGAGAGATGTGTGAGCTGA	1244
QY	1081	TGAAGCAGCAGAGAGAGCAGCTGAAAGAGCAGTACAGTCCAGAGAGATGTGTGAGCTGA	1140
Db	1245	TGAAGCAGCAGAGAGAGCAGCTGAAAGAGCAGTACAGTCCAGAGAGATGTGTGAGCTGA	1304
QY	1141	AGTTCCAGAGACACATTTTCCAAAAGCAGGAGTATTCACCATTCACAGCAGAGATGG	1200
Db	1305	AGTTCCAGAGACACATTTTCCAAAAGCAGGAGTATTCACCATTCACAGCAGAGATGG	1364
QY	1201	AAAGATGATCTAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1260
Db	1365	AAAGATGATCTAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1424
QY	1261	GGGAGAGCAGCAACAGGCGCTGCTTGGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Db	1425	GGGAGAGCAGCAACAGGCGCTGCTTGGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAG	1484
QY	1321	AACCTGGAGGCGCTCAGGTAAGATCCAAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG	1380
Db	1485	AACCTGGAGGCGCTCAGGTAAGATCCAAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG	1544
QY	1381	CAGAGCGCAATGACCTGAAACAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1440
Db	1545	CAGAGCGCAATGACCTGAAACAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1604
QY	1441	TCATGACAGTGGCGCTGAGAGAGGCGCAGAGGCGCTGGGGCTCAAGAGAGAGAGAGAG	1500
Db	1605	TCATGACAGTGGCGCTGAGAGAGGCGCAGAGGCGCTGGGGCTCAAGAGAGAGAGAGAG	1664
QY	1501	CCAGGCTCAGAGAGAGCGCTTGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1560
Db	1665	CCAGGCTCAGAGAGAGCGCTTGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1724
QY	1561	CTGGGCTCAG	1620
Db	1725	CTGGGCTCAG	1784
QY	1621	GAAGGAGCGGAG	1680
Db	1785	GAAGGAGCGGAG	1844

RESULT 8

US-09-976-740-46
Sequence 46, Application US/09976740
Publication No. US20020194633A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Aribal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/976,740
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 09/615,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547

; PRIOR FILING DATE: 1997-11-20

827	ACTTCAGGTGACACTGAAATGCACATTCAGCTGCGAGATGGAAACAGCACAAATGAGCGCAACT	885
841	CCAAAGCTGCGCCAAAGAGAAACAATGGAGTGTGGCTGAGAGGCTTCAAGAAGCTGATGAGCAGT	900
887	CCAAAGCTGCGCCAAAGAGAAACAATGGAGTGTGGCTGAGAGGCTTCAAGAAGCTGATGAGCAGT	946
901	ATGAGCTTGGCGGAGGAGCATATCGACAAGATCTTTCAACAACAAGGAGCTTCAACAGAGCAGC	960
947	ATGAGCTTGGCGGAGGAGCATATCGACAAGATCTTTCAACAACAAGGAGCTTCAACAGAGCAGC	1006
961	TGGTGGATGTCCAAGCTCCAGCAGGCCCAAGGAGATGCTTAAAGGAGGCGAGAGCGGGCACC	1020
1007	TGGTGGATGTCCAAGCTCCAGCAGGCCCAAGGAGATGCTTAAAGGAGGCGAGAGCGGGCACC	1066
1021	AGCGGAGAGAGGATTTTCTCTGAAAGAGGCGAGTGTAGAGTCCCGAGAGATGTGTGAGCTGA	1080
1067	AGCGGAGAGAGGATTTTCTCTGAAAGAGGCGAGTGTAGAGTCCCGAGAGATGTGTGAGCTGA	1126
1081	TGAAGCAGCAAGAGACCCACTTGAAGCAACAAGCTTGCCCTTATACACAGAGAAAGTTTGAGG	1140
1127	TGAAGCAGCAAGAGACCCACTTGAAGCAACAAGCTTGCCCTTATACACAGAGAAAGTTTGAGG	1186
1141	AGTTCCAGAAACACACTTTTCCAAAAGCAGCGAGGTATTACCAACATTCAAGCAGGAGATGG	1200
1187	AGTTCCAGAAACACACTTTTCCAAAAGCAGCGAGGTATTACCAACATTCAAGCAGGAGATGG	1246
1201	AAAAGATGACTAAGAGATCAAGAAAGCTGTGAGAAAGAAAACACCATGTACCGGTCCCGGT	1260
1247	AAAAGATGACTAAGAGATCAAGAAAGCTGTGAGAAAGAAAACACCATGTACCGGTCCCGGT	1306
1261	GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGTAGGAGAAAACACTTCGGGATTAAG	1320
1307	GGGAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGTAGGAGAAAACACTTCGGGATTAAG	1366
1321	AACCTGGAGGGGCTGCAGTAAATAATCCAAACGGCTGGAGAAAGCTGTGCGGCACTGCAGA	1380
1367	AACCTGGAGGGGCTGCAGTAAATAATCCAAACGGCTGGAGAAAGCTGTGCGGCACTGCAGA	1426
1381	CAGAGCGGATGACCTGAAACAAGAGGGTACAGGACCTGAGTGTGTGTGGCCAGAGGCTCCC	1440
1427	CAGAGCGCAATGACCTGAAACAAGAGGGTACAGGACCTGAGTGTGTGTGGCCAGAGGCTCCC	1486
1441	TCATGTACAGTGGCCCTGTAGAGGAGGGCCAGAGGGGCTTGGGGCTCAAGACCCAGACTCCC	1500
1487	TCATGTACAGTGGCCCTGTAGAGGAGGGCCAGAGGGGCTTGGGGCTCAAGACCCAGACTCCC	1546
1501	CCAGGGTCACAGAAGCGCTTGCTTACCCAGAGGACCCGAGCACAAGAAGCATCAGGCCAGA	1560
1547	CCAGGGTCACAGAAGCGCTTGCTTACCCAGAGGACCCGAGCACAAGAAGCATCAGGCCAGA	1606
1561	CTGGGGCTCAGAGGCCCACTTCGCGCAGGGCC	1592
1607	CTGGGGCTCAGAGGCCCACTTCGCGCAGGGCC	1638

RESULT 10

US-10-023-529-46

Sequence 46, Application US/10023529

Publication No. US20020129388A1

GENERAL INFORMATION:

APPLICANT: Lees, Ann M.

APPLICANT: Lees, Robert S.

APPLICANT: Law, Simon W.

APPLICANT: Arjona, Anibal, A.

1. TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING

; TITLE OF INVENTION: ATHEROSCLEROSIS

; FILE REFERENCE: 10797-004001

;; CURRENT APPLICATION NUMBER: US/10/023,529

;; CURRENT FILING DATE: 2001-12-17

;; PRIOR APPLICATION NUMBER: 09/616,289

; PRIOR FILING DATE: 2000-07-14

;; PRIOR APPLICATION NUMBER: US 09/517,849

```

; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-10-671-242-46

```

```
Query Match      30.6%; Score 1439; DB 13; Length 1638;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

1	CAAAAGACGCCAGAGCAAAACCGGAAGCAGGACCCGAGAGCGGGCCGACGCC	60
47	CAAAAGACGCCAGGACAAACCGGAAGCAGGACCCGAGAGCGGGCCGACGCC	106
61	AGCGGGCTCTGCAGTAGAGCAGAGAGGTCCCGGCAGCAGCCAGGCTCCTCGAAGCGGG	120
107	AGCGGGCTCTGCAGTAGAGCAGAGAGGTCCCGGCAGCAGCCAGGCTCCTCGAAGCGGG	166
121	AGGGGGCTCAAGCCAGCAACGGCTCAGTCTGGGGCCCTTCGTGATGTTCTCTGAGGAGCTGA	180
167	AGGGTGTCTCAAGCCAGCAACGGCTCAGTCTGGGGCCCTTCGTGATGTTCTCTGAGGAGCTGA	226
181	GCGCCCAACTGGGAAGACATCTACTGAGCACATCTGTGTGGACAATTAACACAGAGGGGGCCCGG	240
227	GCGCCCAACTGGGAAGACATCTACTGAGCACATCTGTGTGGACAATTAACACAGAGGGGGCCCGG	286
241	GCAGAGATGGGGCACAGGGGTGACCGGGCTGAACCCGGAAGATGCAGAGAAGTCCCGGACCT	300
287	GCAGAGATGGGGCACAGGGGTGACCGGGCTGAACCCGGAAGATGCAGAGAAGTCCCGGACCT	346
301	ATGTGGCAAGGAATGGGGAGCCCTGAAACCAACTCCAGTAGTCTATGGAGAGAAGAAACCTT	360
347	ATGTGGCAAGGAATGGGGAGCCCTGAAACCAACTCCAGTAGTCAATGGAGAGAAGAAACCTT	406
361	CCAGGGGGATCCAAACACAGAGAAGATCCGCACAGATGCAGAGTCCGAGACCCGAGACC	420
407	CCAGGGGGATCCAAACACAGAGAAGATCCGCACAGATGCAGAGTCCGAGACCCGAGACC	466
421	ATCGAAGCCACAGGAGAGAAAGCAAGGGTTTGGGGAAAGAGATCACGTTGCTGA	480
467	ATCGAAGCCACAGGAGAGAAAGCAAGGGTTTGGGGAAAGAGATCACGTTGCTGA	526
481	TGCAGACATTGAATCTCTGAGTACCCACAGAGAGAAGCTGGCTCTCTGTGCAAGAAGT	540
527	TGCAGACATTGAATCTCTGAGTACCCACAGAGAGAAGCTGGCTCTCTGTGCAAGAAGT	586
541	ATGCTGAACCTGCTGGAGGAGCACCGGAATTCAACAGACAGATGAAGCTCTCTACAGAAA	600
587	ATGCTGAACCTGCTGGAGGAGCACCGGAATTCAACAGACAGATGAAGCTCTCTACAGAAA	646
601	AGCAGAGCCAGCTGGTGCAAGAGAAGGACCACTCGCGCGGTGAGCACAGCAAGCCGCTCC	660
647	AGCAGAGCCAGCTGGTGCAAGAGAAGGACCACTCGCGCGGTGAGCACAGCAAGCCGCTCC	706
661	TGGCCCGCAGCAAGCTTTGAGAGCCTATGCGGTGAGCTGCAGCGGCACAAACCGCTCCCTCA	720
707	TGGCCCGCAGCAAGCTTTGAGAGCCTATGCGGTGAGCTGCAGCGGCACAAACCGCTCCCTCA	766
721	AGGAGAGAGGTGTGCACCGGGCCGGAGGAGGAGAGCGCAAGAGAGTGAACCTCGC	780
767	AGGAGAGAGGTGTGCACCGGGCCGGAGGAGGAGAGCGCAAGAGAGTGAACCTCGC	826
781	ACTTCCAGGTGCACACTGAATGCAATTCAGCTGCAGATGGAAACAGCACAAATGAGCGCAACT	840

PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 1638
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1638)
US-10-023-529-46

Query Match 30.6%; Score 1439; DB 14; Length 1638;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CAAAAGAGCCAGGACCAACCGGAAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGAGCC 60
DB 47 CAAAAGAGCCAGGACCAACCGGAAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGAGCC 106
QY 61 AGCGGCTCTGAGTAGAAGAGAGAGGTCCTCGGAGCAGCAGGCTCTCGAAGCCGG 120
DB 107 AGCGGCTCTGAGTAGAAGAGAGAGGTCCTCGGAGCAGCAGGCTCTCGAAGCCGG 166
QY 121 AGCGGCTCAAGCAGACCGGCTCAGTCTGGGCGCTTCTGATGTCCTGAGGAGCTGA 180
DB 167 AGCGGCTCAAGCAGACCGGCTCAGTCTGGGCGCTTCTGATGTCCTGAGGAGCTGA 226
QY 181 GCCGCCAATGGGAAGACATCTAGACACATATCTGTGTGACAAATTAACAGGGGGGCGCG 240
DB 227 GCCGCCAATGGGAAGACATCTAGACACATATCTGTGTGACAAATTAACAGGGGGGCGCG 286
QY 241 GCCAGGATGGGACAGGCTGAGCGGCTGAACCCGAAGTCAGAGAGTCCCGGACCT 300
DB 287 GCCAGGATGGGACAGGCTGAGCGGCTGAACCCGAAGTCAGAGAGTCCCGGACCT 346
QY 301 ATGTGCAAGGAATGGGAGCCTGAACCAACTCCAGTCTATGAGAGAGGAAACCT 360
DB 347 ATGTGCAAGGAATGGGAGCCTGAACCAACTCCAGTCTATGAGAGAGGAAACCT 406
QY 361 CCAAGGGGATCAACACAGAGAGATCGGAGAGTCCGAGAGTCCGAGAGCCGAGCC 420
DB 407 CCAAGGGGATCAACACAGAGAGATCGGAGAGTCCGAGAGTCCGAGAGCCGAGCC 466
QY 421 ATCCAGGCGCACAGGAG 480
DB 467 ATCCAGGCGCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 526
QY 481 TGCAAGCATTTGAATCTCTGAGTACCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 527 TGCAAGCATTTGAATCTCTGAGTACCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 586
QY 541 ATGCTGAATCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 587 ATGCTGAATCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646
QY 601 AGCAGAGCCAGTCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 647 AGCAGAGCCAGTCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
QY 661 TGCCCGCGCAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 707 TGCCCGCGCAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
QY 721 AGGAAGAGTGTGACGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 767 AGGAAGAGTGTGACGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826

QY 781 ACTTCCAGGTGACACTGAATGAATTCAGCTGAGATGGAACAGCAGCAATGAGCGCACT 840
DB 827 ACTTCCAGGTGACACTGAATGAATTCAGCTGAGATGGAACAGCAGCAATGAGCGCACT 886
QY 841 CCAAGCTCGCCAGAGAGAAATGAGCTGCTGAGAGCTCAAGAGCTGATTGAGCAGT 900
DB 887 CCAAGCTCGCCAGAGAGAAATGAGCTGCTGAGAGCTCAAGAGCTGATTGAGCAGT 946
QY 901 ATGAGCTCGCGAGGAGCATATCGAACAAGCTTCAACAAGAGGACTCAACAGCAGC 960
DB 947 ATGAGCTCGCGAGGAGCATATCGAACAAGCTTCAACAAGAGGACTCAACAGCAGC 1006
QY 961 TGGTGGATGCAAGCTCCAGAGCCCGAGAGAGTCTTAAGGAGGAGAGAGAGCGGACC 1020
DB 1007 TGGTGGATGCAAGCTCCAGAGCCCGAGAGAGTCTTAAGGAGGAGAGAGAGCGGACC 1066
QY 1021 AGCGGAGAGAGAGATTTCTCTCTGAAAGAGGAGCTAGAGTCCAGAGAGATCTGTAGCTGA 1080
DB 1067 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGCTAGAGTCCAGAGAGATCTGTAGCTGA 1126
QY 1081 TGAAGAGCAG 1140
DB 1127 TGAAGAGCAG 1186
QY 1141 AGTTCCAGAGACACACTTTCCAAAGAGAGAGAGTATTACACCAATTCAGAGAGAGATGG 1200
DB 1187 AGTTCCAGAGACACACTTTCCAAAGAGAGAGAGTATTACACCAATTCAGAGAGAGATGG 1246
QY 1201 AAAAGATGACTAAGAGATCAAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1247 AAAAGATGACTAAGAGATCAAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
QY 1261 GGGAGAGCAGACCAAGCGCTCTGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1307 GGGAGAGCAGACCAAGCGCTCTGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366
QY 1321 AACTGGAGGGCTGAGAGTAAATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1367 AACTGGAGGGCTGAGAGTAAATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1426
QY 1381 CAGAGCCAGATGACTGAAACAGAGAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1427 CAGAGCCAGATGACTGAAACAGAGAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1486
QY 1441 TCAGTCAAGTGGGCTGAG 1500
DB 1487 TCAGTCAAGTGGGCTGAG 1546
QY 1501 CCAGGGTCAAG 1560
DB 1547 CCAGGGTCAAG 1606
QY 1561 CTGGGCTCAAG 1592
DB 1607 CTGGGCTCAAG 1638

RESULT 11

US-10-023-523-46
; Sequence 46, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17

661 TGGCCCGCAGCAAGCTTGAGAGCCATAIGCCGTGAGGCTGCAAGCGCTCCCTCA 720

US-09-976-740-51
; Sequence 51, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Artona, Anibal A.

; Sequence 51, Application US/09976740
 ; Publication No. US20020194633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arlona, Anibal A.

;/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;/ TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;/ TITLE OF INVENTION: ATHEROSCLEROSIS
;/ FILE REFERENCE: 10797-004001
;/ CURRENT APPLICATION NUMBER: US/09/976,740
;/ CURRENT FILING DATE: 2001-10-12
;/ PRIOR APPLICATION NUMBER: 09/616,289
;/ PRIOR FILING DATE: 2000-07-14
;/ PRIOR APPLICATION NUMBER: US 08/979,608
;/ PRIOR FILING DATE: 1997-11-26
;/ PRIOR APPLICATION NUMBER: US 60/031,930
;/ PRIOR FILING DATE: 1996-11-27
;/ PRIOR APPLICATION NUMBER: US 60/048,547
;/ PRIOR FILING DATE: 1997-06-03
;/ NUMBER OF SEQ ID NOS: 53
;/ SOFTWARE: FastSeq for Windows Version 4.0
;/ SEQ ID NO 51
;/ LENGTH: 22255
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ US-09-976-740-51

Query Match 18.0%; Score 844; DB 9; Length 22255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 6; Indels 9; Gaps 8;

QY	1300	AGAAACAGTCCGGGTAAGAACTGGAGGCGCTGCGAGTAAATCCAAAGGCTGGAGA	1359
DB	16637	AGAAACAGTCCGGGTAAGAACTGGAGGCGCTGCGAGTAAATCCAAAGGCTGGAGA	16696
QY	1360	AGCTGTGCGGGCACTGACAGACAGAGCGCAATGACCTGAAACAGAGGGGTACAGGACCTGA	1419
DB	16697	AGCTGTGCGGGCACTGACAGACAGAGCGCAATGACCTGAAACAGAGGGGTACAGGACCTGA	16756
QY	1420	GTGCTGTGGCGAGGGCTCCCTCACTGACAGTGGCGCTGAGAGGAGGCGCCAGAGGGCGCTG	1479
DB	16757	GTGCTGTGGCGAGGGCTCCCTCACTGACAGTGGCGCTGAGAGGAGGCGCCAGAGGGCGCTG	16816
QY	1480	GGGCTCAAGCACCAGCTCCCGAGGCTCACAGAGCGCTTGTACACAGAGGACACCGA	1539
DB	16817	GGGCTCAAGCACCAGCTCCCGAGGCTCACAGAGCGCTTGTACACAGAGGACACCGA	16876
QY	1540	GCACAGAAGCATCAGGCCAGCTGGGCTTCAAGAGCCCACTCCCGAGGGCTTAGAGAG	1599
DB	16877	GCACAGAAGCATCAGGCCAGCTGGGCTTCAAGAGCCCACTCCCGAGGGCTTAGAGAG	16936
QY	1600	CCTGGTGTGGCTCATGCTGGAGAGGAGCGGCGCCAGCCAGCGCTGGCCCATAAAG	1659
DB	16937	CCTGGTGTGGCTCATGCTGGAGAGGAGCGGCGGCGCCAGCCAGCGCTGGCCCATAAAG	16996
QY	1660	GCTCCCATGCTGAGCAGCCCATTTGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTG	1719
DB	16997	GCTCCCATGCTGAGCAGCCCATTTGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTG	17055
QY	1720	GCATTTGCAATTTTGGATTTTGGTGTAGTTTACGTACATAGGCGCATTTTGCAGGGCC	1779
DB	17056	GCATTTGCAATTTTGGATTTTGGTGTAGTTTACGTACATAGGCGCATTTTGCAGGGCC	17115
QY	1780	TTGCAATTCATTTTACCTGTAAGGTACAGTGGGCTTGGCAATGGGGATGGGGTGTGT	1839
DB	17116	TTGCAATTCATTTTACCTGTAAGGTACAGTGGGCTTGGCAATGGGGATGGGGTGTGT	17175
QY	1840	ACAGATGAAGTCAAGTGGTGTGTGTGAGCTGAAAGTCTTACAGAGGGGCTGTACATCTG	1899
DB	17176	ACAGATGAAGTCAAGTGGTGTGTGTGAGCTGAAAGTCTTACAGAGGGGCTGTACATCTG	17235
QY	1900	AGCTGCCATCAGTGGTGGCAGAGTGACTTGGAGCATTTCTGTGCTGATTTGAGGC	1959
DB	17236	AGCTGCCATCAGTGGTGGCAGAGTGACTTGGAGCATTTCTGTGCTGATTTGAGGC	17295
QY	1960	TCAGACCCCTCCCTGCGCTTTCAGAGCTCAAAACAGTAATACCAAGGCTTGTGACTGC	2019
DB	17296	TCAGACCCCTCCCTGCGCTTTCAGAGCTCAAAACAGTAATACCAAGGCTTGTGACTGC	17354

QY	2020	ATTGTCTTGTGACAGGGTGTCTTGGTCAAGTCAAGCCCTCTAGCTCT-TGGAGGC	2078
DB	17355	ATTGTCTTGTGACAGGGTGTCTTGGTCAAGTCAAGCCCTCTAGCTCTCTGAGGC	17414
QY	2079	TCCTTTGATTTCTAGACCTTGGAAAGGTCTCCTTAGCAGAGCCCTGGCAGGCGCTCA	2138
DB	17415	TCCTTTGATTTCTAGACCTTGGAAAGGTCTCCTTAGCAGAGCCCTGGCAGGCGCTCA	17474
QY	2139	GAGCT-GGGATTTCTGCTGCTGGAACAAAGGACCTGGAGAAATGTTTTGCGTGGGATGATG	2197
DB	17475	GAGCTGGGGATTTGCTGCTGGAACAAAGGACCTGGAGAAATGTTTTGCGTGGGATGATG	17534
QY	2198	TGCTGGTCAAGAGCCCTTGGGCAATCGCTTCCCTGCGCTTGGTAGTGCAGGACCG	2257
DB	17535	TGCTGGTCAAGAGCCCTTGGGCAATCGCTTCCCTGCGCTTGGTAGTGCAGGACCG	17594
QY	2258	CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGCTTGCACAAATGA	2317
DB	17595	CCAATGATGCTTCTCAGTAGCTTATCATTCACAGGTGCTCTCTAGCTTGCACAAATGA	17654
QY	2318	TTGCAAGAGATCACCCAAAGGATTTATTCGAAAGGTGTTTTTCTTTATTTCTTTTC	2377
DB	17655	TTGCAAGAGATCACCCAAAGGATTTATTCGAAAGGTGTTTTTCTTTATTTCTTTTC	17714
QY	2378	--TTTTTTTTTTTTTTTTTTTTTTTTTGTGACATGACAGTGTGTTGATTGAGGACC	2435
DB	17715	TTTTTTTTTTTTTTTTTTTTTTTTTGTGACATGACAGTGTGTTGATTGAGGACC	17774
QY	2436	TTCAAGGAAAGGATGCTGTACAGTGTGCTGGTGGCTGGCCCTCAGTGTCCAC	2495
DB	17775	TTCAAGGAAAGGATGCTGTACAGTGTGCTGGTGGCTGGCCCTCAGTGTCCAC	17834
QY	2496	CTCCTTCAACACCCAGCTTGGCTCTTGGCATCTTGATCTGAGTGTCTGTTTGGTG	2555
DB	17835	CTCCTTCAACACCCAGCTTGGCTCTTGGCATCTTGATCTGAGTGTCTGTTTGGTG	17894
QY	2556	AGATCAGGTGTGTTGGTAAAGAAAGAGGAAAGGGCTTCTGATGGCTTGGCACAAGCTT	2615
DB	17895	AGATCAGGTGTGTTGGTAAAGAAAGAGGAAAGGGCTTCTGATGGCTTGGCACAAGCTT	17954
QY	2616	ACCTGTGGGTTTCACTGCTGAGAGGCCACACACAGTCCCATCAGACTGTCTCCATGCA	2675
DB	17955	ACCTGTGGGTTTCACTGCTGAGAGGCCACACACAGTCCCATCAGACTGTCTCCATGCA	18014
QY	2676	GCAGTGTGGGTCCTCAGCTGCTTCTTGGCTTCTGATGGTGTCTGCTTCTGCTCTG	2735
DB	18015	GCAGTGTGGGTCCTCAGCTGCTTCTTGGCTTCTGATGGTGTCTGCTTCTGCTCTG	18074
QY	2736	CCCCACCCCCACATGTGCAATCCTCAAGATTTGCTGATTTCTATTTCTTGGCACCTCC	2795
DB	18075	CCCCACCCCCACATGTGCAATCCTCAAGATTTGCTGATTTCTATTTCTTGGCACCTCC	18134
QY	2796	CTGCTGTCTTGGGATTTCTTCTTCTGCTGGG-GCCCATAGCTGTGTCTAACA	2854
DB	18135	CTGCTGTCTTGGGATTTCTTCTTCTGCTGGGAGCCCATAGCTGTGTCTAACA	18194
QY	2855	GCTAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAATCCATAAAATGGCTGCAGAC	2914
DB	18195	GCTAAGAAATGAATTTGAATTTGACTGGGCCCCAGAAATCCATAAAATGGCTGCAGAC	18254
QY	2915	AGTTGTTTCTGTGTCCTGTTCTACCCCACTCCAGTACATAACTACTATGTACTGTAG	2974
DB	18255	AGTTGTTTCTGTGTCCTGTTCTACCCCACTCCAGTACATAACTACTATGTACTGTAG	18314
QY	2975	AGCCATCTATATCTGAATGTTCTGCTGTGCAAACTTCCAGGGTATTAGCCAGTGT	3034
DB	18315	AGCCATCTATATCTGAATGTTCTGCTGTGCAAACTTCCAGGGTATTAGCCAGTGT	18374
QY	3035	TGTGCCAAGCAGTTTTCGGGGAACACAGATGACTCAGACCAAGATGGATGGTTA	3094
DB	18375	TGTGCCAAGCAGTTTTCGGGGAACACAGATGACTCAGACCAAGATGGATGGTTA	18434

QY 3095 GGGCTTGTCTTCTGCTGTTTCTTTTGA-CTAG-TGATTTGCTCTGAGGTCCCTTCAT 3152
Db 18435 GGGCTTGTCTTCTGCTGTTTCTTTTGAAGCTAGTTCATTTCTCTGCGAGGTCCCTTCAT 18494
QY 3153 CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA 3212
Db 18495 CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA 18554
QY 3213 CAAAGAATGTTAAGTACTGAATCATGTGTGACTGAGACAGAGATGGCAATGAATGGCA 3272
Db 18555 CAAAGAATGTTAAGTACTGAATCATGTGTGACTGAGACAGAGATGGCAATGAATGGCA 18614
QY 3273 CACCAATTTCTCTCTCTGCCCCAGGCGAGGTACCACTGATCTGTCATCAGAGTTGCCCTG 3332
Db 18615 CACCAATTTCTCTCTCTGCCCCAGGCGAGGTACCACTGATCTGTCATCAGAGTTGCCCTG 18674
QY 3333 CTATTCCTGCTGTATCTCTTCAATCTAGTGCCTTCAAGCAGCTGTGTGAGTTGAGA 3392
Db 18675 CTATTCCTGCTGTATCTCTTCAATCTAGTGCCTTCAAGCAGCTGTGTGAGTTGAGA 18734
QY 3393 TCTCTGCAATCTGCTGCTGAGTACTGCTGCTGTGAGTCTTCCCATGACCTTTTC 3452
Db 18735 TCTCTGCAATCTGCTGCTGAGTACTGCTGCTGTGAGTCTTCCCATGACCTTTTC 18794
QY 3453 TTCCCTTTTGAATCCCTCT 3471
Db 18795 TTCCCTTTTGAATCCCTCT 18813

RESULT 14
US-10-671-242-51
; Sequence 51, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastaSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-671-242-51

Query Match 18.0%; Score 844; DB 13; Length 22255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 6; Indels 9; Gaps 8;

QY 1300 AGAAAAAGTCCGGGATAAAGAACTGGAGGCGCTGAGGTAAAAATCCAAAGCGCTGGAGA 1359
Db 16637 AGAAAAAGTCCGGGATAAAGAACTGGAGGCGCTGAGGTAAAAATCCAAAGCGCTGGAGA 16696
QY 1360 AGCTGTGCGGGGCACTGCGAGACAGAGCGGCAATGACCTGAACAGAGGGGTACAGGACCTGA 1419
Db 16697 AGCTGTGCGGGGCACTGCGAGACAGAGCGGCAATGACCTGAACAGAGGGGTACAGGACCTGA 16756

QY 1420 GTGCTGTGTCAGGAGCTCCCTCACTGACAGTGGCCTGAGAGGCGCAGAGGGCCCTG 1479
Db 16757 GTGCTGTGTCAGGAGCTCCCTCACTGACAGTGGCCTGAGAGGCGCAGAGGGCCCTG 16816
QY 1480 GGGCTCAAGCACCAGCCTCCCCCAGGGTCAAGAAGCGCCTTGTCTACCCAGGAGCACCGA 1539
Db 16817 GGGCTCAAGCACCAGCCTCCCCCAGGGTCAAGAAGCGCCTTGTCTACCCAGGAGCACCGA 16876
QY 1540 GCACAGAAGCATCAGGCCAGACTGGGCTCAAGAGCCCACTCCGCCAGGCGCTAGAGAG 1599
Db 16877 GCACAGAAGCATCAGGCCAGACTGGGCTCAAGAGCCCACTCCGCCAGGCGCTAGAGAG 16936
QY 1600 CCTGGTGTGGGTTCATGCTGGGAAGGAGGCGGAGCCAGCCAGGCTGGCCCCATAAAG 1659
Db 16937 CCTGGTGTGGGTTCATGCTGGGAAGGAGGCGGAGCCAGGCTGGCCCCATAAAG 16996
QY 1660 GCTCCCATGCTGAGCAGCCCAATTTGCTGAAGCAGAGATGTTCTTGAAGCTGGCTGGCATCTG 1719
Db 16997 GCTCCCATGCTGAGCAGCCCAATTTGCTGAAGCAGAGATGTTCTTGAAGCTGGCTGGCATCTG 17055
QY 1720 GCACCTTCAATTTTGGATTTTGTGGGTGAGTTTACGTACATAGGCAATTTTCAAGGCC 1779
Db 17056 GCACCTTCAATTTTGGATTTTGTGGGTGAGTTTACGTACATAGGCAATTTTCAAGGCC 17115
QY 1780 TTGCAAAATGCAATTTATACCTGTAAAGTGTACAGTGGGCTTCATTTGGGGATGGGGTGTGT 1839
Db 17116 TTGCAAAATGCAATTTATACCTGTAAAGTGTACAGTGGGCTTCATTTGGGGATGGGGTGTGT 17175
QY 1840 ACAGATCAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAAGGCGGTGTCATCTGT 1899
Db 17176 ACAGATCAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAAGGCGGTGTCATCTGT 17235
QY 1900 AGCTGCCATCAAGTGGAGTGGCAGAGTCAATTTGAGCAATTTCTGTCTGTGATTTGAGGC 1959
Db 17236 AGCTGCCATCAAGTGGAGTGGCAGAGTCAATTTGAGCAATTTCTGTCTGTGATTTGAGGC 17295
QY 1960 TCAGACCCCTCCCTGCTTTCAGAGCTCAAAACAGAGTAATACCAAGTCTTGAAGTGC 2019
Db 17296 TCAGACCCCTCCCTGCTTTCAGAGCTCAAAACAGAGTAATACCAAGTCTTGAAGTGC 17354
QY 2020 ATTTGTCTTGTGAGCAGGCTTGTCTGTGAGTCAAGGCGCTCTAGTCTCT-TGGAGGC 2078
Db 17355 ATTTGTCTTGTGAGCAGGCTTGTCTGTGAGTCAAGGCGCTCTAGTCTCT-TGGAGGC 17414
QY 2079 TCCTTTGATTTCTTAGACCTTGGAAAAAGGTCTCCCTAGGAGAGCCCTTGGCAGGCGGTCA 2138
Db 17415 TCCTTTGATTTCTTAGACCTTGGAAAAAGGTCTCCCTAGGAGAGCCCTTGGCAGGCGGTCA 17474
QY 2139 GAGCT-GGGATTTCTGCTGGACAGGCGACCTGGAGATGTTTTTTCCTGGGATGATG 2197
Db 17475 GAGCTGGGATTTGCTGCTGGACAGGCGACCTGGAGATGTTTTTTCCTGGGATGATG 17534
QY 2198 TGCTGTGTCAGGAGCCCTTGGGATCGCTTCCCTGCGCTTTTGGTAGTGCAGGACGAG 2257
Db 17535 TGCTGTGTCAGGAGCCCTTGGGATCGCTTCCCTGCGCTTTTGGTAGTGCAGGACGAG 17594
QY 2258 CCAATGATGCTTCTCAGTACGCTTATCATTACAGGTGCGCTCTAGCTGCAATGA 2317
Db 17595 CCAATGATGCTTCTCAGTACGCTTATCATTACAGGTGCGCTCTAGCTGCAATGA 17654
QY 2318 TTGCAAGAGATCACCCAAAAGGATTTTCTGAAGGTGTTTTTCTTTATTTCTTTTTC 2377
Db 17655 TTGCAAGAGATCACCCAAAAGGATTTTCTGAAGGTGTTTTTCTTTATTTCTTTTTC 17714
QY 2378 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCACATGACAGTGTGTTATTTAGAGAC 2435
Db 17715 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCACATGACAGTGTGTTATTTAGAGAC 17774
QY 2436 TTCCAAAGAAAAGGATGCTGACAGTGGTGGCTGGCTGGCTCCAGTGTCCAC 2495
Db 17775 TTCCAAAGAAAAGGATGCTGACAGTGGTGGCTGGCTGGCTCCAGTGTCCAC 17834

Qy	1900	AGCTGCCATCACAGTGAGTGTGGCAGAAGTGACTTGAGCANNTTCTCTGTCTGATTTGAGGC	1959
Db	17236	AGCTGCCATCACAGTGAGTGTGGCAGAAGTGACTTGAGCANNTTCTCTGTCTGATTTGAGGC	17295
Qy	1960	TCAGACCCCTCCCTGGCCCTTTAGAGCTCAAAACAAAGTAATAACACCAAGGCTCTTGACTGC	2019
Db	17296	TCAGACCCCTCCCTGGCCCTTTAGAGCTCAAGCAAGTAATAACACCCAGGCTCTTGACTGC	17354
Qy	2020	ATTTGTCTTTGTGAGCAGGCGTTCCTTTGGTCAGCTCAGGCCCTCTCTAGTGCTCTTGAGGC	2078
Db	17355	ATTTGTCTTTGTGAGCAGGCGTTCCTTTGGTCAGCTCAGGCCCTCTCTAGTGCTCTTGAGGC	17414
Qy	2079	TCCTTTGATTTCTAGACCTGGAAAAAGTGTCCCTTAGCAGAGCCCTGGCAGGGCGCTCA	2138
Db	17415	TCCTTTGATTTCTAGACCTGGAAAAAGTGTCCCTTAGCAGAGCCCTGGCAGGGCGCTCA	17474
Qy	2139	GAGCT--GGGATTTTCCTCCCTCGAAACAAAGGCACTCGAGAAATGTTTTTGCCTGGGATGATG	2197
Db	17475	GAGCTGGGATTTTCCTCCCTCGAAACAAAGGCACTCGAGAAATGTTTTTGCCTGGGATGATG	17534
Qy	2198	TGCTGTGTCAGGAGCCCTTTGGGGATCGCTTCCCTGCCCCCTTTGGTAGTGCCAGGACCAAG	2257
Db	17535	TGCTGTGTCAGGAGCCCTTTGGGGATCGCTTCCCTGCCCCCTTTGGTAGTGCCAGGACCAAG	17594
Qy	2258	CCAATGATGCTTCTCAGTAGCCCTTATCATTACAGGTGCCTCTCTAGGCTCGCACAAATGA	2317
Db	17595	CCAATGATGCTTCTCAGTAGCCCTTATCATTACAGGTGCCTCTCTAGGCTCGCACAAATGA	17654
Qy	2318	TTGACAAGAGATCACCCAAAGGATTAATTTCTGAAGTGTTTTTTCTTTATTTCTTTTTTC	2377
Db	17655	TTGACAAGAGATCACCCAAAGGATTAATTTCTGAAGTGTTTTTTCTTTATTTCTTTTTTC	17714
Qy	2378	--TTTTTTTTTTTTCTTTTTCTTTTTTTTTTTTTTTTGGCATGACAGCTGTTGTATTCAGGACC	2435
Db	17715	TTTTTTTTTTTTTTTTCTTTTTCTTTTTTTTTTTTTTTTGGCATGACAGCTGTTGTATTCAGGACC	17774
Qy	2436	TTCCAAGGAAAGGGATGCTTACCAAGTGGTGCTCGGGTGCCCTGCCCTCCAGTGTCCAC	2495
Db	17775	TTCCAAGGAAAGGGATGCTTACCAAGTGGTGCTCGGGTGCCCTGCCCTCCAGTGTCCAC	17834
Qy	2496	CTCGTTCACACACCCACTGGCTCTTCGCCATCTTGATGCTGAGGTTTCCGTGTTGGTG	2555
Db	17835	CTCGTTCACACACCCACTGGCTCTTCGCCATCTTGATGCTGAGGTTTCCGTGTTGGTG	17894
Qy	2556	AGATCAGTGTGTTGTGGTAAAGAAAGGAAAGGGCTTCTGATGCTCTTGCCACAAAGCTT	2615
Db	17895	AGATCAGTGTGTTGTGGTAAAGAAAGGAAAGGGCTTCTGATGCTCTTGCCACAAAGCTT	17954
Qy	2616	ACCTGTGGGTTTCAGTCTCTGAGAGGCCACACACAGTTCACATCAGACTGTCTCATGCA	2675
Db	17955	ACCTGTGGGTTTCAGTCTCTGAGAGGCCACACACAGTTCACATCAGACTGTCTCATGCA	18014
Qy	2676	GCAGTTGCTGGGTCCCATGTCCAGCTGCTCTTTGGCTTCATGGGTGTTTTTCTGTTCTCTG	2735
Db	18015	GCAGTTGCTGGGTCCCATGTCCAGCTGCTCTTTGGCTTCATGGGTGTTTTTCTGTTCTCTG	18074
Qy	2736	CCCCACCCCCACATGCGAAATTCCTCAGATTGTCTGATTTCTATTTCCCTGGCACCTCC	2795
Db	18075	CCCCACCCCCACATGCGAAATTCCTCAGATTGTCTGATTTCTATTTCCCTGGCACCTCC	18134
Qy	2796	CTGCCTGTCTCTGGGGATTTCTACTTCTTCTCTGTGGG--GCCCATAGCTGTGTGCTTAACA	2854
Db	18135	CTGCCTGTCTCTGGGGATTTCTACTTCTTCTCTGTGGG--GCCCATAGCTGTGTGCTTAACA	18194
Qy	2855	GGTAAGAAATGAAATTTGAACATAATGACTGGGCCCCAGAAATCCATAAATGGCTGCAGAC	2914
Db	18195	GGTAAGAAATGAAATTTGAACATAATGACTGGGCCCCAGAAATCCATAAATGGCTGCAGAC	18254
Qy	2915	AGTTGTTTCTGTGCTCTGTTCTACCCCACTCCAGTACATAAATCTACTGTACTGTGTAG	2974
Db	18255	AGTTGTTTCTGTGCTCTGTTCTACCCCACTCCAGTACATAAATCTACTGTACTGTGTAG	18314
Qy	2975	AGCCATCTATATGCTGAATGTTTCTGCTGTGCAAACTTGCACGGGTATATAGCCAGTGTT	3034

Db	18315	AGCATTCTATATGCTGAATGTTCTGCTGTGGAACTTGCAGGGTATTTAGCCAGTGT	18374
Qy	3035	TGTGCCAAGCAGTTTTTCGGGGCAACAGAGAATGACTCAGACCAGATGGATAGGATGGTTA	3094
Db	18375	TGTGCCAAGCAGTTTTTCGGGGCAACAGAGAATGACTCAGACCAGATGGATAGGATGGTTA	18434
Qy	3095	GGCGTTTGCTTCTTGCTGTTTTCTTTTGAA-CTAG-TCATTGCTCCTGCAGGTCCTTTCAT	3152
Db	18435	GGCGTTTGCTTCTTGCTGTTTTCTTTGAAGCTAGTTCAATTGTCCTGCAGGTCCTTTCAT	18494
Qy	3153	CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAATGTGGTTCA	3212
Db	18495	CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAATGTGGTTCA	18554
Qy	3213	CAAGAAGTGTTAAGTACTGAATCATGTGTGACTGAGACCCAGAGATGGCAAAATGAATGGCA	3272
Db	18555	CAAGAAGTGTTAAGTACTGAATCATGTGTGACTGAGACCCAGAGATGGCAAAATGAATGGCA	18614
Qy	3273	CACATTTTCCTTCTCTGCGCCAGGGCAGGTACGACTGATCTGCATCAGAGTGCCTG	3332
Db	18615	CACATTTTCCTTCTCTGCGCCAGGGCAGGTACGACTGATCTGCATCAGAGTGCCTG	18674
Qy	3333	CTATTCTCTGGTGATACCTTTCACATCTAGTGCCCTCAAGCAGCTGTGTGAGTGTGAGA	3392
Db	18675	CTATTCTCTGGTGATACCTTTCACATCTAGTGCCCTCAAGCAGCTGTGTGAGTGTGAGA	18734
Qy	3393	TCTCTGCCATCTCTGACTCAGATCTGCTGCTGTGAGTGTGTTCCTCCATGACCTTTTC	3452
Db	18735	TCTCTGCCATCTCTGACTCAGATCTGCTGCTGTGAGTGTGTTCCTCCATGACCTTTTC	18794
Qy	3453	TTCCCTCTTTGAATCCCTCT	3471
Db	18795	TTCCCTCTTTGAATCCCTCT	18813

RESULT 16

```

US-10-023-523-51
; Sequence 51, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004C01
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51

```

Query Match 18.0%; Score 844; DB 14; Length 2255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 6; Indels 9;

8

Db	18195	GGTAAGAAATGAATGAACATAATGTGACTGGGCCCCAGAAATCCATAAAATGGCTGCAGAC	182554
Qy	2915	AGTTGTTTCTGTGCTCTGCTTTACCCCCCACTCCAGTACATACTACTACTGTGTAG	2974
Db	18255	AGTTGTTTCTGTGCTCTGCTTTACCCCCCACTCCAGTACATACTACTACTGTGTAG	18314
Qy	2975	AGCCATTATATGCTGGAATGTTCTGCTGTTGCAAACTTGCACGGTATATGACGAGTT	3034
Db	18315	AGCCATTATATGCTGGAATGTTCTGCTGTTGCAAACTTGCACGGTATATGACGAGTT	18374
Qy	3035	TGTGCCAAGCAGTTTTGCGGGCAACACAGAAATGACTCAGACCAAGATGCATAGGATGGTTA	3094
Db	18375	TGTGCCAAGCAGTTTTCTGGGCAACACAGAAATGACTCAGACCAAGATGCATAGGATGGTTA	18434
Qy	3095	GGGCTTTTGCTTCTTGCTGTTTTTCTTTTGAA-CTAG-TCATTGCTCTGCAGGTCCTTCAT	3152
Db	18435	GGGCTTTTGCTTCTTGCTGTTTTTCTTTGGAAGCTAGTTCAATTGCTCTGCAGGTCCTTCAT	18494
Qy	3153	CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA	3212
Db	18495	CTTCCATACCTAGCCCACTCTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCA	18554
Qy	3213	CAAGAAGTTTAAGTACTGAATCATGTGTGACTTGAGACCAGAGATGGCAAAATGAATGGCA	3272
Db	18555	CAAGAAGTTTAAGTACTGAATCATGTGTGACTTGAGACCAGAGATGGCAAAATGAATGGCA	18614
Qy	3273	CACCAATTCCTCTCTCTGCCCCAGGGCAGGTACACATGATCTGCAATCAGATTTGCCTG	3332
Db	18615	CACCAATTCCTCTCTCTGCCCCAGGGCAGGTACACATGATCTGCAATCAGATTTGCCTG	18674
Qy	3333	CTATTCTCTGGTGATCCTTTCACATCTAGGTGCCCTCAAGCAGCTGTGTGAGTGTTGAGA	3392
Db	18675	CTATTCTCTGGTGATCCTTTCACATCTAGGTGCCCTCAAGCAGCTGTGTGAGTGTTGAGA	18734
Qy	3393	TCTCTGCCATCTCTGGCTGAGATACCTGCTGCTGTGAAAGTGTTTCCCATGACCTTTTTTC	3452
Db	18735	TCTCTGCCATCTCTGGCTGAGATACCTGCTGCTGTGAAAGTGTTTCCCATGACCTTTTTTC	18794
Qy	3453	TTCCCTTTTGAATCCCTCT	3471
Db	18795	TTCCCTTTTGAATCCCTCT	18813

RESULT 18
US-10-040-739-869
; Sequence 869, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-167

Query Match 7.4%; Score 347; DB 9; Length 1926;
Best Local Similarity 99.5%; Pred. No. 1.4e-156;
Matches 637; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
QY 2834 GCCCATAGCTGTTGCTTAACAGGTAAAGAAATGAAGTAACTTACCTGGGCCCCAGAA 2893
Db 84 GCCCATAGCTGTTGCTTAACAGGTAAAGAAATGAAGTAACTTACCTGGGCCCCAGAA 143
QY 2894 ATCCATAAAATGGCTGCAGACAGTTGTTCTGTGCTCTTCTTACCCCACTCCAGTACA 2953
Db 144 ATCCATAAAATGGCTGCAGACAGTTGTTCTGTGCTCTTCTTACCCCACTCCAGTACA 203
QY 2954 TAACTACTATGTAAGTGTAGAGCCATTTCTATATGCTGAATGTTCTGCTGTTGCAAACTT 3013
Db 204 TAACTACTATGTAAGTGTAGAGCCATTTCTATATGCTGAATGTTCTGCTGTTGCAAACTT 263
QY 3014 GCCAGGGTATTAGCCAGTGTGTTGTCGCAAGCAGTTTTCGGGGACACAGAAATGACTCAGA 3073
Db 264 GCCAGGGTATTAGCCAGTGTGTTGTCGCAAGCAGTTTTCGGGGACACAGAAATGACTCAGA 323
QY 3074 CCAAGATGATAGATGTTAGGGCTTTGCTTCTGCTGTTTCTTTTCTTTTCTTTTCTTTT 3131
Db 324 CCAAGATGATAGATGTTAGGGCTTTGCTTCTGCTGTTTCTTTTCTTTTCTTTTCTTTT 383
QY 3132 TTGCTCTGAGGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3191
Db 384 TTGCTCTGAGGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 443
QY 3192 TCTCTCAGATAAGTGTGTTTCAAAAGATGTTAAAGTAACTGATGCTGATGCTGAGACC 3251
Db 444 TCTCTCAGATAAGTGTGTTTCAAAAGATGTTAAAGTAACTGATGCTGATGCTGAGACC 503
QY 3252 AGAGATGGCAATGAATGACACACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3311
Db 504 AGAGATGGCAATGAATGACACACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 563
QY 3312 GATCTGCATCAGATGCTGCTTATCTCTGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCT 3371
Db 564 GATCTGCATCAGATGCTGCTTATCTCTGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCT 623
QY 3372 GCAGCTGTGAGTGTGATGCTCTGCAATCTCTGCTGATGATGCTGCTGCTGCTGCTGAA 3431
Db 624 GCAGCTGTGAGTGTGATGCTCTGCAATCTCTGCTGATGATGCTGCTGCTGCTGCTGAA 683
QY 3432 GTGTTTCCCATGACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3471
Db 684 GTGTTTCCCATGACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 723
```

RESULT 20

US-918-995-3048/c
; Sequence 3048, Application US/09918995
; Publication No. US20030073623A1

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3048
; LENGTH: 507

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(507)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-3048

Query Match 5.8%; Score 273; DB 10; Length 507;
Best Local Similarity 100.0%; Pred. No. 8.1e-121;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3575 TTTGCTAGGCGAGAGAGGTGCGAGGAAGAGAGAGATGACCTTCCCTTTGAAGAG 3634
Db 331 TTTGCTAGGCGAGAGAGGTGCGAGGAAGAGAGAGATGACCTTCCCTTTGAAGAG 272
QY 3635 AGGGGAAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACTCTTAA 3694
Db 271 AGGGGAAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACTCTTAA 212
QY 3695 TGCTGTGAGAGGACTGAGAGTCTGATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3754
Db 211 TGCTGTGAGAGGACTGAGAGTCTGATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 152
QY 3755 GCAGGTGCCCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCAGGCAAGGTGCGACC 3814
Db 151 GCAGGTGCCCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCAGGCAAGGTGCGACC 92
QY 3815 TTTGCTCAGAGTCCATCCATGCTAGACCTTC 3847
Db 91 TTTGCTCAGAGTCCATCCATGCTAGACCTTC 59
```

RESULT 21

US-09-918-995-26710
; Sequence 26710, Application US/09918995
; Publication No. US20030073623A1

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26710
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(481)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26710

Query Match 5.7%; Score 268; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.1e-118;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3575 TTTGCTAGGCGAGAGAGGTGCGAGGAAGAGAGAGATGACCTTCCCTTTGAAGAG 3634
Db 214 TTTGCTAGGCGAGAGAGGTGCGAGGAAGAGAGAGATGACCTTCCCTTTGAAGAG 273
QY 3635 AGGGGAAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACTCTTAA 3694
Db 274 AGGGGAAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACTCTTAA 333
QY 3695 TGCTGTGAGAGGACTGAGAGTCTGATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3754
Db 334 TGCTGTGAGAGGACTGAGAGTCTGATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 393
```

QY 3755 GCAGGTGGCCCCAGCTGGGTTTCCAGGTCAGGAATGGGGCCCGAGCAAGGTGCAGCC 3814
Db 394 GCAGGTGGCCCCAGCTGGGTTTCCAGGTCAGGAATGGGGCCCGAGCAAGGTGCAGCC 453
QY 3815 TTGTCTCAGAGTCATCATCTCTAGA 3842
Db 454 TTGTCTCAGAGTCATCATCTCTAGA 481

RESULT 22

US-09-918-995-11546/c
; Sequence 11546, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11546
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-11546

Query Match 5.6%; Score 264; DB 10; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.9e-116; Indels 0; Gaps 0;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3208 GTTCAAGAAGATGTTAAGTACTGAATCATGTGTGACTGAGACCCAGAGATGCAATGAA 3267
Db 419 GTTCAAGAAGATGTTAAGTACTGAATCATGTGTGACTGAGACCCAGAGATGCAATGAA 360
QY 3268 TGGCACACCATTTCTCTCTGCTGCCCCAGGGAGGTACCACTGATCTGCATCAGAGTT 3327
Db 359 TGGCACACCATTTCTCTCTGCTGCCCCAGGGAGGTACCACTGATCTGCATCAGAGTT 300
QY 3328 GCTGTGATTTCTGTGTGATCTTACATCTAGGTGCCCTCAAGCAGCTGTGTGAGTGT 3387
Db 299 GCTGTGATTTCTGTGTGATCTTACATCTAGGTGCCCTCAAGCAGCTGTGTGAGTGT 240
QY 3388 TGAGATCTCTGCCATCTCTGGCTGAGATCTCTGTCTGTGGAAGTGTTCCTCCATGACCT 3447
Db 239 TGAGATCTCTGCCATCTCTGGCTGAGATCTCTGTCTGTGGAAGTGTTCCTCCATGACCT 180
QY 3448 TTTTCTTCCCTTTGAATCCCTCT 3471
Db 179 TTTTCTTCCCTTTGAATCCCTCT 156

RESULT 23

US-09-920-300A-1038
; Sequence 1038, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038

; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1038
Query Match 5.0%; Score 235; DB 9; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.9e-102; Indels 0; Gaps 0;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 CTGAGAGCTTCATGTCTCCACCAGATTCTGAGAGGTGTGAGAGCTTCAGGACACTTTTATTTT 4223
Db 1 CTGAGAGCTTCATGTCTCCACCAGATTCTGAGAGGTGTGAGAGCTTCAGGACACTTTTATTTT 60
QY 4224 GTTGTGTTTTCATGAGGTTCGAGCCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 4283
Db 61 GTTGTGTTTTCATGAGGTTCGAGCCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 120
QY 4284 GACTGTTTATTTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCCATCAAGTGGCCCTG 4343
Db 121 GACTGTTTATTTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCCATCAAGTGGCCCTG 180
QY 4344 AGGCTGACCCGGAGGCGGCTGAGCTGCTGTGTCTCCAGGGGAGGGCCAAAG 4398
Db 181 AGGCTGACCCGGAGGCGGCTGAGCTGCTGTGTCTCCAGGGGAGGGCCAAAG 235

RESULT 24

US-10-033-528-1038
; Sequence 1038, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1038

Query Match 5.0%; Score 235; DB 14; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.9e-102; Indels 0; Gaps 0;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 CTGAGAGCTTCATGTCTCCACCAGATTCTGAGAGGTGTGAGAGCTTCAGGACACTTTTATTTT 4223
Db 1 CTGAGAGCTTCATGTCTCCACCAGATTCTGAGAGGTGTGAGAGCTTCAGGACACTTTTATTTT 60
QY 4224 GTTGTGTTTTCATGAGGTTCGAGCCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 4283
Db 61 GTTGTGTTTTCATGAGGTTCGAGCCATCGGCTGAGCTCAGGACACTTTCTGTAGGA 120
QY 4284 GACTGTTTATTTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCCATCAAGTGGCCCTG 4343
Db 121 GACTGTTTATTTCTGTAAGATGGTTATTTAAACCTCTCTCCACCCCATCAAGTGGCCCTG 180
QY 4344 AGGCTGACCCGGAGGCGGCTGAGCTGCTGTGTCTCCAGGGGAGGGCCAAAG 4398
Db 181 AGGCTGACCCGGAGGCGGCTGAGCTGCTGTGTCTCCAGGGGAGGGCCAAAG 235

RESULT 25

US-10-099-926-1038
; Sequence 1038, Application US/10099926
; Publication No. US20030166064A1


```

; NAME/KEY: misc_feature
; LOCATION: 226, 245
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-1375

Query Match      4.8%   Score 225;   DB 15;   Length 250;
Best Local Similarity 100.0%; Pred. No. 1.3e-97;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4194  AGGTGTCAGCAGCACTTTTTTTTTTTATTTGTTGTTTGTTCCTCAGAGTTATCGGACC 4253
Db      225  AGGTGTCAGCAGCACTTTTTTTTTTTATTTGTTGTTTGTTCCTCAGAGTTATCGGACC 166

Qy      4254  ATGGGCTGAGCTCAGGCACCTTCTGTAGGAGACGTGTTATTTCCTAAAGATGTTATTTA 4313
Db      165  ATGGGCTGAGCTCAGGCACCTTCTGTAGGAGACTGTTATTTCCTAAAGATGTTATTTA 106

Qy      4314  ACCCTCTCTCCACCCCATCAGCGTGGCCCTCAGGGCTCACCCGGAGGCCAGTGGAGCTGCC 4373
Db      105  ACCCTCTCTCCACCCCATCAGCGTGGCCCTCAGGGCTCACCCGGAGGCCAGTGGAGCTGCC 46

Qy      4374  TGTGTGTCACGGGGGAGGGCCAAAGGCTGTGAGCTGATTCCTCA 4418
Db      45    TGTGTGTCACGGGGGAGGGCCAAAGGCTGTGAGCTGATTCCTCA 1

```

```

RESULT 29
US-09-833-381-1334
; Sequence 1334, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1334
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(492)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1334

```

	Query Match.	4.6%;	Score 218;	DB 9;	Length 492;
	Best Local Similarity	100.0%;	Pred. No. 2.9e-94;		
	Matches 218;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2834	GCCCATAGCTGTTGTTCTTAACAGGTAAAGAAATGAACTATTGACTGGGCCCCAGAA	2893		
DB	201	GCCCATAGCTGTTGTTCTTAACAGGTAAAGAAATGAACTATTGACTGGGCCCCAGAA	260		
QY	2894	ATCCATAAAATGGCTGCAGACAGTTGTTCTGTGTCCTGTTCACCCCACTCCAGTACA	2953		
DB	261	ATCCATAAAATGGCTGCAGACAGTTGTTCTGTGTCCTGTTCACCCCACTCCAGTACA	320		
QY	2954	TAACTACTATGTAAGTGTAGAGCCATTCATATAGCTGAATGTTCTGCTGTGCACAACTT	3013		
DB	321	TAACTACTATGTAAGTGTAGAGCCATTCATATAGCTGAATGTTCTGCTGTGCACAACTT	380		
QY	3014	GCCAGGTATTAGCCAGTGTGTTGCCAAGCAGTTTTC	3051		
DB	381	GCCAGGTATTAGCCAGTGTGTTGCCAAGCAGTTTTC	418		

RESULT 30
US-09-962-436-33

```

; Sequence 33, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
;
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
; US-09-962-436-33

```

	Query Match	3.5%;	Score 166;	DB 9;	Length 441;
	Best Local Similarity	100.0%;	Pred. No. 3.8e-69;		
	Matches 166;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3306	ACCACGTGATCTGCATCAGAGTTCGCTGCTATTCCTCGGTGATCCCTTCACATCTAGGTGC	3365		
Db	43	ACCACGTGATCTGCATCAGAGTTCGCTGCTATTCCTCGGTGATCCCTTCACATCTAGGTGC	102		
Qy	3366	CCTCAAGCAGCTGTGTGAGTGTGGAGATCTCTGCCATCTCTGGCTGAGATCTGCTCTCC	3425		
Db	103	CCTCAAGCAGCTGTGTGAGTGTGGAGATCTCTGCCATCTCTGGCTGAGATCTGCTCTCC	162		
Qy	3426	TGTGAAGTGTTCCTCCATGACCTTTTCTCCCTTTGAATCCCTCT	3471		
Db	163	TGTGAAGTGTTCCTCCATGACCTTTTCTCCCTTTGAATCCCTCT	208		

```

RESULT 31
US-10-085-783A-11181
/ Sequence 11181, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liaw, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ CURRENT FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 11181
/ LENGTH: 209
/ TYPE: DNA
/ ORGANISM: Human
US-10-085-783A-11181

```

Query Match	3.5%;	Score 165;	DB 13;	Length 209;
Best Local Similarity	100.0%;	Pred. No. 1.2e-68;		
Matches 165;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	4346	GGCTGACCCGGAGGCCAGTGGAGCTCCCTGGTGTCCACGGGGGAGGCCAAGCCCTCTG	4405	
Db	1	GGCTGACCCGGAGGCCAGTGGAGCTCCCTGGTGTCCACGGGGGAGGCCAAGCCCTCTG	60	


```
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10263
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10263

Query Match          3.0%; Score 143; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4555 CCGGCCACTCAGACTTGTGAAGCTGCACTGGGAGCTTTTGTCTCTCTTTGGGTATT 4614
Db 142 CCGGCCACTCAGACTTGTGAAGCTGCACTGGGAGCTTTTGTCTCTCTTTGGGTATT 201
QY 4615 CACAACAGCCAGGAGCTTGTATTTGATGATTTTAAACACACATTAATAAAGAGCTGTT 4674
Db 202 CACAACAGCCAGGAGCTTGTATTTGATGATTTTAAACACACATTAATAAAGAGCTGTT 261
QY 4675 GCCTTAAAAAATAAAAAAAAAAAAAA 4697
Db 262 GCCTTAAAAAATAAAAAAAAAAAAAA 284

RESULT 36
US-09-764-869-467
; Sequence 467, Application US/09764869
; Publication No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-467

Query Match          3.0%; Score 140; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.4e-56;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 243
Db 88 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 147
QY 244 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 303
Db 148 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 207
QY 304 TGGCAAGGAATGGGAGCCT 323
Db 208 TGGCAAGGAATGGGAGCCT 227

RESULT 37
US-09-764-869-467
; Sequence 467, Application US/09764869
; Publication No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-467

Query Match          3.0%; Score 140; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.4e-56;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 243
Db 88 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 147
QY 244 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 303
Db 148 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 207
QY 304 TGGCAAGGAATGGGAGCCT 323
Db 208 TGGCAAGGAATGGGAGCCT 227

RESULT 38
US-10-023-523-17.rnbp
; Sequence 467, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-504-467

Query Match          3.0%; Score 140; DB 15; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.4e-56;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 243
Db 88 GCCAACTGGAAGACATACACTGAGCACAATCTGTGTGACAAATACCAAGGGGGCCCCGGCG 147
QY 244 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 303
Db 148 AGGATGGGCGACAGGTGAGCCGGTGAACCCGGAAGATGCAGAGAAGTCCCGGACCTATG 207
QY 304 TGGCAAGGAATGGGAGCCT 323
Db 208 TGGCAAGGAATGGGAGCCT 227

RESULT 39
US-10-227-577-467
; Sequence 467, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 467

LENGTH: 405

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (377)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (398)

OTHER INFORMATION: n equals a,t,g, or c

US-10-227-577-467

Query Match 3.0%; Score 140; DB 16; Length 405;

Best Local Similarity 100.0%; Pred. No. 1.4e-56;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 GCCAACTGGAAGACATCTAGACACATCTGTGTGACAAATACCGAGGGGGGCGCGG 243

DB 88 GCCAACTGGAAGACATCTAGACACATCTGTGTGACAAATACCGAGGGGGGCGCGG 147

QY 244 AGGATGGGCGACAGGCTGAGCGGCTGAACCCGAGATGCAGAGAGTCCCGGACCTATG 303

DB 148 AGGATGGGCGACAGGCTGAGCGGCTGAACCCGAGATGCAGAGAGTCCCGGACCTATG 207

QY 304 TGGCAAGGAATGGGAGCGCT 323

DB 208 TGGCAAGGAATGGGAGCGCT 227

RESULT 39

US-09-918-995-14235

Sequence 14235, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 14235

LENGTH: 416

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(416)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-14235

Query Match

Best Local Similarity 2.7%; Score 129; DB 10; Length 416;

Matches 229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4450 CACAGAGTGGTTCACCCGAGGACAGCCAGGACCTGCTCTTGGCCCTTCTGGGGGA 4509

DB 1 CACAGAGTGGTTCACCCGAGGACAGCCAGGACCTGCTCTTGGCCCTTCTGGGGGA 60

QY 4510 AAGAGTGGCTTCTGTCCTGTAACTGCTTCTTATGSCCAACCCGCGCACTCAGAC 4569

DB 61 AAGAGTGGCTTCTGTCCTGTAACTGCTTCTTATGSCCAACCCGCGCACTCAGAC 120

QY 4570 TTGTTTGAAGTGCACCTGGGAGCTTTTGTCTCTTTGGGTATTCACACAGCCAGGGA 4629

DB 121 TTGTTTGAAGTGCACCTGGGAGCTTTTGTCTCTTTGGGTATTCACACAGCCAGGGA 180

QY 4630 CTTGATTTTGATGATTTTAAACACACATTAATAAAGAGTCTGTGCTTA 4680

DB 181 CTTGATTTTGATGATTTTAAACACACATTAATAAAGAGTCTGTGCTTA 231

RESULT 40

US-09-918-995-11012

Sequence 11012, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 11012

LENGTH: 547

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(547)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-11012

Query Match

Best Local Similarity 2.7%; Score 127; DB 10; Length 547;

Matches 227; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4452 CAGAGTGGTTCACCCGAGGACAGCCAGGACCTGCTCTTGGCCCTTCTGGGGGAAA 4511

DB 32 CAGAGTGGTTCACCCGAGGACAGCCAGGACCTGCTCTTGGCCCTTCTGGGGGAA 91

QY 4512 GGAGTGGCTTCTGTCCTGTAACTGCTTTTCTTATGGCCCAACCCGCGCACTCAGACTT 4571

DB 92 GGAGTGGCTTCTGTCCTGTAACTGCTTTCTTATGGCCCAACCCGCGCACTCAGACTT 151

QY 4572 GTTTGAAGTGCACCTGGCAGCTTTTGTCTCTTGGGTATTCACACAGCCAGGACT 4631

DB 152 GTTTGAAGTGCACCTGGCAGCTTTTGTCTCTTGGGTATTCACACAGCCAGGACT 211

QY 4632 TGATTTGATGATTTTAAACACACATTAATAAAGAGTCTGTGCTTA 4680

DB 212 TGATTTGATGATTTTAAACACACATTAATAAAGAGTCTGTGCTTA 260

RESULT 41

US-10-060-036-2444

Sequence 2444, Application US/10060036

Publication No. US20030073144A1

GENERAL INFORMATION:

APPLICANT: Benson, Darin R.

APPLICANT: Kalos, Michael D.

APPLICANT: Lodes, Michael J.

APPLICANT: Persing, David H.

APPLICANT: Hepler, William T.

APPLICANT: Jiang, Yugu

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.566

CURRENT APPLICATION NUMBER: US/10/060,036

CURRENT FILING DATE: 2002-01-30

NUMBER OF SEQ ID NOS: 4560

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2444

LENGTH: 264

TYPE: DNA

ORGANISM: Homo sapiens
US-10-060-036-2444

Query Match
Best Local Similarity 2.3%; Score 108; DB 15; Length 264;
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2966 ACTGTGATAGGCAATCTATATGCTGAATGTTCTGCTGTTGCAAACTTGCCAGGGTATTA 3025
DB 1 ACTGTGATAGGCAATCTATATGCTGAATGTTCTGCTGTTGCAAACTTGCCAGGGTATTA 60

QY 3026 GCAGTGTGTTGCGCAAGAGTTTCGGGGACAAACAGATGACTCAGACCAAGATGGATA 3085
DB 61 GCAGTGTGTTGCGCAAGAGTTTCGGGGACAAACAGATGACTCAGACCAAGATGGATA 120

QY 3086 GGATGTTAGGCTTGTCTGCTGCTGCTGTTTCTTTTGA 3124
DB 121 GGATGTTAGGCTTGTCTGCTGCTGCTGTTTCTTTTGA 159

RESULT 42
US-09-962-055-14
; Sequence 14, Application US/09962055
; Patent No. US2002005203A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-Nov-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...1731
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-962-055-14

Query Match
1.3%; Score 59; DB 9; Length 4722;

Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1350 CGGTGGAGAAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGT 1408
DB 1456 CGGTGGAGAAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 43
US-09-976-740-14
; Sequence 14, Application US/09976740
; Publication No. US2002019463A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-09-976-740-14

Query Match
1.3%; Score 59; DB 9; Length 4722;

Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1350 CGGTGGAGAAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGT 1408
DB 1456 CGGTGGAGAAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGT 1514

RESULT 44
US-10-671-242-14
; Sequence 14, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930

;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)...(1731)
US-10-671-242-14

Query Match 1.3%; Score 59; DB 13; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 45
US-10-023-529-14
;; Sequence 14, Application US/10023529
;; Publication No. US20020129388A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; TITLE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,529
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)...(1731)
US-10-023-529-14

Query Match 1.3%; Score 59; DB 14; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 46
US-10-023-523-14
;; Sequence 14, Application US/10023523
;; Publication No. US20020152485A1
;; GENERAL INFORMATION:

;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; TITLE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,523
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)...(1731)
US-10-023-523-14

Query Match 1.3%; Score 59; DB 14; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1350 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1408
Db 1456 CGGCTGGAGAGCTGTGCCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 47
US-10-616-187-14
;; Sequence 14, Application US/10616187
;; Publication No. US20040013668A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; TITLE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/616,187
;; CURRENT FILING DATE: 2003-07-09
;; PRIOR FILING DATE: 2003-07-09
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 4722
;; TYPE: DNA
;; ORGANISM: Oryctolagus cuniculus
;; FEATURE:
;; NAME/KEY: CDS

```
; LOCATION: (61)...(1731)
US-10-616-187-14

Query Match
  1.3%; Score 59; DB 15; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.6e-17; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 0;

QY 1350 CGCGTCGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1408
    |||||
Db 1456 CGCGTCGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 48
US-10-027-632-289698/c
; Sequence 289698, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289698
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289698

Query Match
  0.9%; Score 44; DB 13; Length 522;
Best Local Similarity 100.0%; Pred. No. 3.1e-10; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 0;

QY 4289 TTATTTCTGTAAGATGGTTATTAAACCTCTCCACCCCATCA 4332
    |||||
Db 407 TTATTTCTGTAAGATGGTTATTAAACCTCTCCACCCCATCA 364

RESULT 50
US-09-876-143-1454
; Sequence 1454, Application US/09876143
; Publication No. US20040081958A1
; GENERAL INFORMATION:
; APPLICANT: Infigen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: PEISTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; FILE REFERENCE: 028040-0202
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1454
; LENGTH: 937
; TYPE: DNA
; ORGANISM: Bovine
; NAME/KEY: misc feature
; LOCATION: (1)...(937)
; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1454

Query Match
  0.8%; Score 37; DB 12; Length 937;
Best Local Similarity 100.0%; Pred. No. 7.1e-07; Indels 0; Gaps 0;
Matches 37; Conservative 0; Mismatches 0;

QY 267 GCTCAACCCGGAAGATGCAGAGAGTCCCGGACCTATG 303
    |||||
Db 198 GCTCAACCCGGAAGATGCAGAGAGTCCCGGACCTATG 234

RESULT 51
US-10-221-714A-430/c
; Sequence 430, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
```

```
; LOCATION: (61)...(1731)
US-10-616-187-14

Query Match
  1.3%; Score 59; DB 15; Length 4722;
Best Local Similarity 100.0%; Pred. No. 1.6e-17; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 0;

QY 1350 CGCGTCGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1408
    |||||
Db 1456 CGCGTCGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGGGT 1514

RESULT 48
US-10-027-632-289698/c
; Sequence 289698, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289698
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289698

Query Match
  0.9%; Score 44; DB 13; Length 522;
Best Local Similarity 100.0%; Pred. No. 3.1e-10; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 0;

QY 4289 TTATTTCTGTAAGATGGTTATTAAACCTCTCCACCCCATCA 4332
    |||||
Db 407 TTATTTCTGTAAGATGGTTATTAAACCTCTCCACCCCATCA 364

RESULT 50
US-09-876-143-1454
; Sequence 1454, Application US/09876143
; Publication No. US20040081958A1
; GENERAL INFORMATION:
; APPLICANT: Infigen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: PEISTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; FILE REFERENCE: 028040-0202
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1454
; LENGTH: 937
; TYPE: DNA
; ORGANISM: Bovine
; NAME/KEY: misc feature
; LOCATION: (1)...(937)
; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1454

Query Match
  0.8%; Score 37; DB 12; Length 937;
Best Local Similarity 100.0%; Pred. No. 7.1e-07; Indels 0; Gaps 0;
Matches 37; Conservative 0; Mismatches 0;

QY 267 GCTCAACCCGGAAGATGCAGAGAGTCCCGGACCTATG 303
    |||||
Db 198 GCTCAACCCGGAAGATGCAGAGAGTCCCGGACCTATG 234

RESULT 51
US-10-221-714A-430/c
; Sequence 430, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
```

```

RESULT 55:
US-09-960-352-2223/c
/ Sequence 2223, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21.(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 2223
/ LENGTH: 415
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ FEATURE:

```

RESULT 53
US-10-424-599-123984
; Sequence 123984, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

Query Match 0.6%; Score 30; DB 13; Length 479;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 60

```

US-09-764-891-9923
; Sequence 9923, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Roset et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9923
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-9923

```

Query Match 0.6%; Score 30; DB 10; Length 2696;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 61
US-10-027-632-112265/c
; Sequence 112265, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:

```

: APPLICANT: Wang, David G.
:
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
:
: TITLE OF INVENTION: Polymorphisms in the Human Genome
:
: FILE REFERENCE: 108827.129
:
: CURRENT APPLICATION NUMBER: US/10/027.632
:

```

```

/ CURRENT_FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 112265
/ LENGTH: 2752

```

; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-112265

```
Query Match      0.6%; Score 30; DB 13; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 2378 TTTTTTTTTTCTTTCTTTTTT 2407
2507 TTTTTTTTTTCTTTCTTTTTT 2478
Db 2507 TTTTTTTTTTCTTTCTTTTTT 2478

RESULT 62
US-10-027-632-112265/c
; Sequence 112265, Application US/10027632
; Publication No. US20030204075A9

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome

```

```

FILE REFERENCE: 106827.123
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 112265

```

; LENGTH: 2752
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-11226

Query Match

Query Match 0.6%; Score 30; DB 16; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 30; Conservative 0; Mismatches 0; Indels

[illegible]

RESULT 63

US-10-312-841-1/c
; Sequence 1, Application US/10312841
: Publication No. US20030186277A1

GENERAL INFORMATION:

APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC

```

; FILE OF INVENTION: DIAGNOSTIC VEH. RECORDS
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2

```

; SEQ ID NO 1

; LENGTH: 3673778

TYPE: DNA

```

; ORGANISM: Artificial sequence
: FEATURE:

```

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

FEATURE:

; NAME/KEY: unsure

; LOCATION: (3294164)

US-10-312-841-1

```
Query Match      0.6%; Score 30; DB 15; Length 3673778;
Best Local Similarity 100.0%; Pred.No. 0.0011;
Matches 30; Conservative 0; Mismatches 0; Indels 0;
```

Qy 2378 TTTTTTTTTTCTTTTCTTTTTTTTTT 2407
 |||||
 Db 2119327 TTTTTTTTTTCTTTTCTTTTTTTTTT 2119

RESULT 64

US-10-424-599-129561
; Sequence 129561, Application US/10424599
; Publication No. US20040031072A1

RESULT 66
US-10-027-632-143303/c
: Sequence 143303. Application US/10027632

Query Match 0.68; Score 29; DB 16; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2368 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2396
Db 360 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 332
RESULT 68
US-10-027-632-143303/c
; Sequence 143303, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143303
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-143303

Query Match 0.6%; Score 29; DB 16; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2368 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2396
Db 360 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 332

RESULT 69
US-10-027-632-261090
; Sequence 261090, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261090
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-261090

Query Match 0.6%; Score 29; DB 13; Length 634;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2379 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2407
Db 333 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 361

RESULT 70
US-10-027-632-261090
; Sequence 261090, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261090
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-261090

Query Match 0.6%; Score 29; DB 16; Length 634;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2379 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2407
Db 333 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 361

RESULT 71
US-10-027-632-117473
; Sequence 117473, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

TYPE: DNA

RESULT 90
US-10-292-798-1393
Sequence 1393, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUWA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1393
LENGTH: 822900
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(822900)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(1068)
FEATURE:
NAME/KEY: CDS
LOCATION: (140545)..(140693)
FEATURE:
NAME/KEY: CDS
LOCATION: (261786)..(261845)
FEATURE:
NAME/KEY: CDS
LOCATION: (273663)..(273702)
FEATURE:
NAME/KEY: CDS
LOCATION: (347633)..(347711)
FEATURE:
NAME/KEY: CDS
LOCATION: (482589)..(482596)
FEATURE:
NAME/KEY: CDS
LOCATION: (534176)..(534210)
FEATURE:
NAME/KEY: CDS
LOCATION: (822485)..(822700)
FEATURE:
NAME/KEY: modified base
LOCATION: (4848)..(4947)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (4966)..(4966)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (17785)..(17884)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (61159)..(61258)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (67605)..(67704)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (74625)..(74724)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (85854)..(85953)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367573)..(367573)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367588)..(367588)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367595)..(367595)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367613)..(367614)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367985)..(367985)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367993)..(367993)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367998)..(367998)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370273)..(370277)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370279)..(370279)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370281)..(370282)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370285)..(370287)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370289)..(370291)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (385643)..(385742)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (409961)..(410060)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (410096)..(410096)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:

```
/ NAME/KEY: modified base
/ LOCATION: (417384)..(417483)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (724960)..(725059)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (726106)..(726205)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (727470)..(727569)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (728849)..(728948)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (730296)..(730395)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (731863)..(731962)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (734124)..(734223)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (734441)..(734441)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (735752)..(735851)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (738576)..(738675)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (740924)..(741023)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (744360)..(744459)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (748430)..(748529)
/ OTHER INFORMATION: a, t, c, g, unknown or other
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (754323)..(754422)
```

Query Match 0.6%; Score 29; DB 16; Length 822900;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2377 CTTTTTTTTTTCTTTTCTTTTCTTTT 2405
DB 247310 CTTTTTTTTTTCTTTTCTTTTCTTTT 247338
```

RESULT 91

```
US-10-418-182-65
/ Sequence 65, Application US/10418182
/ Publication No. US2003028302A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Crea, Roberto
/ TITLE OF INVENTION: UNIVERSAL LIBRARIES FOR IMMUNOGLOBULINS
/ FILE REFERENCE: 1551.2001-001
/ CURRENT APPLICATION NUMBER: US/10/418,182
/ CURRENT FILING DATE: 2003-04-16
/ PRIOR APPLICATION NUMBER: 60/373,558
/ PRIOR FILING DATE: 2002-04-17
/ NUMBER OF SEQ ID NOS: 423
/ SOFTWARE: PastSeq for Windows Version 4.0
/ SEQ ID NO 65
/ LENGTH: 36
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: oligonucleotide
US-10-418-182-65
```

Query Match 0.6%; Score 28; DB 16; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2380 TTTTTTTTTTTCTTTTCTTTTCTTTT 2407
DB 1 TTTTTTTTTTTCTTTTCTTTTCTTTT 28
```

RESULT 92

```
US-10-424-599-70577/c
/ Sequence 70577, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 70577
/ LENGTH: 280
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34741C.1
US-10-424-599-70577
```

Query Match 0.6%; Score 28; DB 13; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2378 TTTTTTTTTTTCTTTTCTTTTCTTTT 2405
DB 68 TTTTTTTTTTTCTTTTCTTTTCTTTT 41
```

RESULT 93

```
US-09-876-143-359
/ Sequence 359, Application US/09876143
/ Publication No. US20040081958A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Infigen, Inc.
/ APPLICANT: EILERTSEN, KENNETH J.
/ APPLICANT: PFISTER-GENSKOW, MARTHA
/ APPLICANT: CHILDS, LYNETTE
/ APPLICANT: FORSYTHE, TODD
/ APPLICANT: BISHOP, MICHAEL D.
/ TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
/ TITLE OF INVENTION: CELLULAR REPROGRAMMING
/ FILE REFERENCE: 028040-0202
/ CURRENT APPLICATION NUMBER: US/09/876,143
```

```

1 / APPLICANT: Steinmann, Kathleen
2 /
3 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
4 /
5 / TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
6 /
7 / TITLE OF INVENTION: THERAPY OF BREAST CANCER
8 /
9 / FILE REFERENCE: MEI-049
10 /
11 / CURRENT APPLICATION NUMBER: US/10/198,846
12 /
13 / CURRENT FILING DATE: 2002-07-18
14 /

```

```

RESULT 97
US-10-027-632-221256/c
; Sequence 221256, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 07:04:24 ; Search time 1723.84 Seconds
(without alignments)
11575.198 Million cell updates/sec

Title: US-10-023-523-17

Perfect score: 4697
Sequence: 1 caaaagcagccaggagca.....ttaaaaaaaaaaaaaaa 4697

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001bs:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4697	100.0	4697	5	Aah26496 Human low
2	4544	96.7	4697	2	Aav32839 Human low
3	1819	38.7	5085	7	Acc46153 Human dit
4	1819	38.7	5085	7	Acc46130 Human dit
5	1589	33.8	2523	4	Abao09008 Human LDL
6	1589	33.8	2523	4	Aah52874 Human pol
7	1439	30.6	1638	5	Aah26501 Human low
8	1294	27.5	1793	7	Abz34807 Coding se
9	1262	26.9	2188	4	Aak51890 Human pol
10	1156	24.6	1793	3	Aaa35218 Human ade
11	1156	24.6	1793	3	Aaa35217 Human ade
12	1156	24.6	1793	3	Aaf21340 Human low
13	1156	24.6	1793	3	Aaf21339 Human low
14	1156	24.6	1793	7	Abz297034 Human nuc
15	1156	24.6	1793	7	Abz297033 Human nuc
16	895	19.1	115756	7	Accl3448 Human dna
17	824	18.0	22255	5	Aah26497 Human low
18	723	15.4	1854	2	Aaq66987 Interleuk
19	686	14.6	5284	5	Aas65026 DNA encod
20	367	7.8	529	2	Aav88391 EST clone
21	347	7.4	1926	6	Abk35029 Human cdn
22	273	5.8	507	8	Ach15836 Human adu
23	268	5.7	481	8	Ach39498 Human foe

C 24	264	5.6	419	8	ACH24334	Ach24334 Human adu
C 25	259	5.5	2626	5	RA565100	DNA encod
C 26	239	5.1	316	3	AAC05668	Human sec
C 27	235	5.0	235	6	ABK45487	CDNA enco
C 28	225	4.8	250	4	AAS25194	Human ova
C 29	225	4.8	370	5	Aah83841	Human ova
C 30	166	3.5	268	2	AAT23888	Human gen
C 31	166	3.5	441	6	ABL64155	Stomach c
C 32	153	3.3	236	9	ADD49735	Human lun
C 33	152	3.2	457	8	ACH49990	Human lun
C 34	150	3.2	473	5	AAS65097	DNA encod
C 35	143	3.0	439	8	ACH23051	Human adu
C 36	140	3.0	405	4	AAS35582	Human car
C 37	140	3.0	405	9	ADE45661	Human car
C 38	129	2.7	416	8	ACH27023	Human adu
C 39	127	2.7	547	8	ACH23800	Human adu
C 40	108	2.3	264	6	ABV97036	Human pan
C 41	71	1.5	245	3	AAA45144	Human sec
C 42	59	1.3	4722	2	AAV32836	Rabbit lo
C 43	59	1.3	4722	5	RAH26491	Rabbit lo
C 44	37	0.8	937	6	ABN74519	Bovine em
C 45	34	0.7	14615	4	RAA46705	Tumour su
C 46	30	0.6	30	6	ABK66221	Human gen
C 47	30	0.6	30	6	ABK66222	Human gen
C 48	30	0.6	405	5	ABV58328	Human pro
C 49	30	0.6	408	8	ACH20077	Human adu
C 50	30	0.6	415	7	ABX37058	Bovine ES
C 51	30	0.6	2696	4	AAAL07235	Human rep
C 52	30	0.6	2696	4	AAAL07233	Human rep
C 53	30	0.6	2696	4	ABL98781	Human tes
C 54	30	0.6	2696	4	ABL98783	Human tes
C 55	29	0.6	197	4	AAK71901	Human imm
C 56	29	0.6	198	4	AAK71898	Human imm
C 57	29	0.6	198	4	AAK71903	Human imm
C 58	29	0.6	419	5	AAF64499	Novel hum
C 59	29	0.6	1239	7	ABZ77145	Human pro
C 60	29	0.6	1795	9	ADB56171	Toxicity-
C 61	29	0.6	1795	9	ADB53743	Primary r
C 62	29	0.6	4156	4	AAK83412	Human imm
C 63	29	0.6	10894	4	AAK36611	Human car
C 64	29	0.6	10894	4	AAK66343	Human imm
C 65	29	0.6	10894	9	ADE47305	Human car
C 66	29	0.6	47573	9	ADA02840	Human MAP
C 67	29	0.6	47573	9	ADB2578	Human MAP
C 68	29	0.6	47573	9	ADC85319	Mouse Map
C 69	29	0.6	136328	6	ABZ35015	Human gen
C 70	29	0.6	349901	9	ADC86940	Human GPC
C 71	28	0.6	252	4	AAAL25533	Human bre
C 72	28	0.6	265	4	AAAL16690	Human bre
C 73	28	0.6	282	4	AAAL07787	Human bre
C 74	28	0.6	333	6	ABN73417	Bovine em
C 75	28	0.6	341	5	ABV04933	Human pro
C 76	28	0.6	342	5	ABV14102	Human pro
C 77	28	0.6	392	5	ABV44044	Human pro
C 78	28	0.6	392	5	ABV35204	Human pro
C 79	28	0.6	482	6	ABV96912	Human pan
C 80	28	0.6	515	4	AAAL25838	Human bre
C 81	28	0.6	776	4	AAAL19003	Human bre
C 82	28	0.6	1706	4	AAK94092	Human ful
C 83	28	0.6	1870	9	ADB80281	Human pro
C 84	28	0.6	2178	5	ABV22311	Human pro
C 85	28	0.6	2178	5	ABV23125	Human pro
C 86	28	0.6	2178	5	ABV28966	Human pro
C 87	28	0.6	2178	5	ABV28140	Human pro
C 88	28	0.6	2888	3	AACT6962	Human ORF
C 89	28	0.6	3001	3	AAH51799	Chromosome
C 90	28	0.6	3197	8	ACF05254	Plasmodiu
C 91	28	0.6	3905	6	ABN83030	Human X2
C 92	28	0.6	8913	7	ABZ09919	Human 5'
C 93	28	0.6	8913	7	ABZ10218	Haematopo
C 94	28	0.6	8913	7	ABZ10072	Haematopo
C 95	28	0.6	11523	4	AAS45468	Chemical
C 96	28	0.6	11523	4	AAS46642	Tumour su

C 97 28 0.6 11523 6 ABL33923 Human imm
 C 98 28 0.6 11523 6 ABK34002 Human DNA
 C 99 28 0.6 11523 6 ABK28324 DNA trans
 C 100 28 0.6 11523 7 ADA20383 Prostate

ALIGNMENTS

RESULT 1

AAH26496 standard; cDNA; 4697 BP.

AC AAH26496;

DT 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 3 (LBP-3) cDNA.

KW Low density lipoprotein binding protein 3; LBP-3; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
 KW ss.

XX Homo sapiens.

Key Location/Qualifiers
 CDS 3..1596
 FT /*tag= a
 FT /partial

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX P-PSDB; AAB82804.

PT New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Example 4; Fig 17; 143pp; English.

The present sequence is that of a partial cDNA encoding novel human low
 density lipoprotein binding protein 3 (LBP-3, see AAB82804). Clones were
 isolated from human foetal brain, liver and aorta cDNA libraries using
 rabbit LBP-3 cDNA as probe. A full-length cDNA clone is given in
 AAB26501, and a genomic DNA sequence is given in AAB26497. LBP-3 nucleic
 acids are among claimed polynucleotides of the invention that encode
 novel polypeptides capable of binding to native and methylated LDL. Also
 claimed are isolated LBP polypeptides, and biologically active fragments
 and analogues of them, as well as expression vectors, cells and methods
 of producing the LBPs. Methods of determining if an animal is at risk for
 atherosclerosis, methods for evaluating an agent for use in treating
 atherosclerosis, and methods for treating a cell having an abnormality in
 structure or metabolism of LBP are claimed. Pharmaceutical compositions
 comprising an LBP polypeptide or nucleic acid, and vaccine compositions,
 are also claimed

XX Sequence 4697 BP; 1097 A; 1197 C; 1290 G; 1113 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 100.0%; Score 4697; DB 5; Length 4697;

XX Matches 4697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAAGCAGCCCAAGCAACCGAAGCAGGACCCAGGAGCCAGGAGCGGCCAGCC 60
 DB 1 CAAAAGCAGCCCAAGCAACCGAAGCAGGACCCAGGAGCCAGGAGCGGCCAGCC 60
 QY 61 AGCGCGCTCTCTGAGTAGAAGCAGAGAGGTCCCGGACAGCAGCAGGCTCTCGGAAGCGG 120
 DB 61 AGCGCGCTCTCTGAGTAGAAGCAGAGAGGTCCCGGACAGCAGCAGGCTCTCGGAAGCGG 120
 QY 121 AGGGGGCTCAAGCCAGAAACGGCTCAGTCTGGGGCCCTTCTGTGATGTCTCTGAGGAGCTGA 180
 DB 121 AGGGGGCTCAAGCCAGAAACGGCTCAGTCTGGGGCCCTTCTGTGATGTCTCTGAGGAGCTGA 180
 QY 181 GCCGCCAACTGGAAGACATCTGAGCACATCTGTGTGGACAATAACAGGGGGGGCCCCG 240
 DB 181 GCCGCCAACTGGAAGACATCTGAGCACATCTGTGTGGACAATAACAGGGGGGGCCCCG 240
 QY 241 GCAGAGATGGGGCACAGGGTGAAGCTGAACCCGGAAGATCGAGAGAGTCCCGGACCT 300
 DB 241 GCAGAGATGGGGCACAGGGTGAAGCTGAACCCGGAAGATCGAGAGAGTCCCGGACCT 300
 QY 301 ATGTGGCAAGATGGGGACCTGAACCACTCCAGTACTCTATGAGAGAGGACCTTCTGAGAGGACCT 360
 DB 301 ATGTGGCAAGATGGGGACCTGAACCACTCCAGTACTCTATGAGAGAGGACCTTCTGAGAGGACCT 360
 QY 361 CCAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGGAGACCGAGACC 420
 DB 361 CCAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCCGGAGACCGAGACC 420
 QY 421 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGGTTTGGGAGAGGAGATCAGTTGCTGA 480
 DB 421 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGGTTTGGGAGAGGAGATCAGTTGCTGA 480
 QY 481 TGCAGACATTGAACTCTCTGAGTACCCACAGAGAGAGTGGCTCTCTGTGCAAGAAGT 540
 DB 481 TGCAGACATTGAACTCTCTGAGTACCCACAGAGAGAGTGGCTCTCTGTGCAAGAAGT 540
 QY 541 ATGTGAACTGCTGGAGGACCCGGAAATTCACAGAGCAGATGAAGTCTCTCAGAAAA 600
 DB 541 ATGTGAACTGCTGGAGGACCCGGAAATTCACAGAGCAGATGAAGTCTCTCAGAAAA 600
 QY 601 AGCAGAGCAGCTGGTCCAGAGAGAGAGACCACTGCGCGGTGAGCAGCAGCAAGCGCTCC 660
 DB 601 AGCAGAGCAGCTGGTCCAGAGAGAGAGACCACTGCGCGGTGAGCAGCAGCAAGCGCTCC 660
 QY 661 TGGCCCGCAGCAAGCTTTCAGAGCCTATGCGGTGAGCTGCGCGGCACAAACCGCTCCCTCA 720
 DB 661 TGGCCCGCAGCAAGCTTTCAGAGCCTATGCGGTGAGCTGCGCGGCACAAACCGCTCCCTCA 720
 QY 721 AGGAAGAGTGTGTCAGCGGGCCCGGAGGAGGAGAGAGAGAGAGAGAGTACCTCGC 780
 DB 721 AGGAAGAGTGTGTCAGCGGGCCCGGAGGAGGAGAGAGAGAGAGAGAGTACCTCGC 780
 QY 781 ACTTCCAGTGCACACTGAATGACATTTCAGTTCAGATGGAACAGCACAATGAGCGCAACT 840
 DB 781 ACTTCCAGTGCACACTGAATGACATTTCAGTTCAGATGGAACAGCACAATGAGCGCAACT 840
 QY 841 CCAAGCTGCGCCAAAGAGAACATGAGCTGGCTGAGAGGCTCAGAGAGCTGATTGAGCAGT 900
 DB 841 CCAAGCTGCGCCAAAGAGAACATGAGCTGGCTGAGAGGCTCAGAGAGCTGATTGAGCAGT 900
 QY 901 ATGAGCTGCGGAGGAGCATATCCAAAGTCTTCAAAACAAGGACTCAACAGCAGC 960
 DB 901 ATGAGCTGCGGAGGAGCATATCCAAAGTCTTCAAAACAAGGACTCAACAGCAGC 960
 QY 961 TGGTGGATGCCAAGCTCCAGCAGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1020
 DB 961 TGGTGGATGCCAAGCTCCAGCAGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1020
 QY 1021 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGTGAGCTGA 1080
 DB 1021 AGCGGAGAGAGATTTCTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGTGAGCTGA 1080
 QY 1081 TGAAGCAG 1140

1081 TGAGCAGCAAGAGACCCACCTTGAAGCAACAGCTTGGCCCTATACAGAGAAAGTTTGAGG 1140
1141 AGTTCCAGAACACACTTTTCAAAAAGCAGCGAGGTATTACACCACTTCAAGCAGGAGATGG 1200
1141 AGTTCCAGAACACACTTTTCAAAAAGCAGCGAGGTATTACACCACTTCAAGCAGGAGATGG 1200
1201 AAAAGATGATCAAGAGATCAAGAGCTGGAGAAAGAAACCAACATGATACCGGTCCCGGT 1260
1201 AAAAGATGATCAAGAGATCAAGAGCTGGAGAAAGAAACCAACATGATACCGGTCCCGGT 1260
1261 GGGAGAGCAGCAACAGAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATAAAG 1320
1261 GGGAGAGCAGCAACAGAGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATAAAG 1320
1321 AACTGAGGGCTGCAAGGTAAATAACCAACGCTGGAGAGCTGTCCGGGCACTGCAGA 1380
1321 AACTGAGGGCTGCAAGGTAAATAACCAACGCTGGAGAGCTGTCCGGGCACTGCAGA 1380
1381 CAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCC 1440
1381 CAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCC 1440
1441 TCACTGACAGTGGCCCTGAGAGAGGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCC 1500
1441 TCACTGACAGTGGCCCTGAGAGAGGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCC 1500
1501 CAGGGTCAAGAGCGCCCTTGTACCCAGAGAGCAACCGAGCAAGAGCATCAGGCCAGA 1560
1501 CAGGGTCAAGAGCGCCCTTGTACCCAGAGAGCAACCGAGCAAGAGCATCAGGCCAGA 1560
1561 CTGGGCTCAAGAGCGCCCTTGTACCCAGAGAGCTGTGGGCTGTGGGTCACTGCTGG 1620
1561 CTGGGCTCAAGAGCGCCCTTGTACCCAGAGAGCTGTGGGCTGTGGGTCACTGCTGG 1620
1621 GAAAGGAGGGGAGCGCCAGCGAGGCTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCA 1680
1621 GAAAGGAGGGGAGCGCCAGCGAGGCTGGCCCATAAAGGCTCCCATGCTGAGCAGGCCA 1680
1681 TTGCTGAGCAGGATGTTCTTGAAGCTGGCTGGCATCTGGCATTTGCAATTTGATTTT 1740
1681 TTGCTGAGCAGGATGTTCTTGAAGCTGGCTGGCATCTGGCATTTGCAATTTGATTTT 1740
1741 GTGGGTCAAGTTTACGTACATAGGCAATTTTCAAGGCTTGCATAATGCAATTTATACCTG 1800
1741 GTGGGTCAAGTTTACGTACATAGGCAATTTTCAAGGCTTGCATAATGCAATTTATACCTG 1800
1801 TAAAGTACAGTGGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAAGTGGCTTG 1860
1801 TAAAGTACAGTGGGCTTGCATTTGGGATGGGGTGTGTACAGATGAAGTCAAGTGGCTTG 1860
1861 TCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGTAGCTGCCATCAGAGTGAAGTTG 1920
1861 TCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGTAGCTGCCATCAGAGTGAAGTTG 1920
1921 GCAGAGTGAAGTGAAGTCTTCTGTGTGATTTGAGGCTCAGACCCCTCCCTGCGCTTT 1980
1921 GCAGAGTGAAGTGAAGTCTTCTGTGTGATTTGAGGCTCAGACCCCTCCCTGCGCTTT 1980
1981 CAGAGCTCAAAACAAGTAAACCAAGCTTTGACTGCAATTTGCTGTGAGCAGGGCT 2040
1981 CAGAGCTCAAAACAAGTAAACCAAGCTTTGACTGCAATTTGCTGTGAGCAGGGCT 2040
2041 TGCCTTGGTCAAGTCAAGGCTTCTAGCTGCTTGGAGGCTCCCTTGAATCTCTAGACCTGG 2100
2041 TGCCTTGGTCAAGTCAAGGCTTCTAGCTGCTTGGAGGCTCCCTTGAATCTCTAGACCTGG 2100
2101 AAAAGGTGCTCCAGCAGAGCCCTGGCAGGCGCTCAGAGCTGGGATTTCTGCGCTTGA 2160
2101 AAAAGGTGCTCCAGCAGAGCCCTGGCAGGCGCTCAGAGCTGGGATTTCTGCGCTTGA 2160
2161 ACAAGGACCTGGAGATGTTTTTTCGTGGATGATGTGCTGGTCAAGAGGCCCTTGGGC 2220

2161 ACAAGGACCTGGAGATGTTTTTTCGTGGATGATGCTGGTCAAGAGCCCTTGGGC 2220
2221 ATCCCTTCCCTGCGCTTTGGTAGTGCAGAGCCAGGCCAATGATGCTTCTCAGTAGCCT 2280
2221 ATCCCTTCCCTGCGCTTTGGTAGTGCAGAGCCAGGCCAATGATGCTTCTCAGTAGCCT 2280
2281 TATCATTCACAGTGCCTCTCTAGCCTGCACAAATGATGACAAAGATGATCAACCAAGGA 2340
2281 TATCATTCACAGTGCCTCTCTAGCCTGCACAAATGATGACAAAGATGATCAACCAAGGA 2340
2341 TATTTCTGAAGGTGTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2400
2341 TATTTCTGAAGGTGTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2400
2401 TTTTTTGCATGACAGTGTGTTGATTGAGGACCTTCCAAAGGAAAAGGATGCTGTACC 2460
2401 TTTTTTGCATGACAGTGTGTTGATTGAGGACCTTCCAAAGGAAAAGGATGCTGTACC 2460
2461 AGTGGTGGCTGGGTGCTGGCTCCAGTGTCCACCTCTTCCACCCACCTTGGCTCC 2520
2461 AGTGGTGGCTGGGTGCTGGCTCCAGTGTCCACCTCTTCCACCCACCTTGGCTCC 2520
2521 TTTGCCATCTTGATGCTGAGGTTTTCTGTTTGGTGGATCAGGTTGTTTGGTAAAGA 2580
2521 TTTGCCATCTTGATGCTGAGGTTTTCTGTTTGGTGGATCAGGTTGTTTGGTAAAGA 2580
2581 AAGAAAAGGGCTTCTGATGCTTGGCAAGCTTACCTGGGTTTTCAGTCTGAGAGG 2640
2581 AAGAAAAGGGCTTCTGATGCTTGGCAAGCTTACCTGGGTTTTCAGTCTGAGAGG 2640
2641 CCACACAGTTCCTCATCAGCACTGTCTCATGAGAGTGTGGGTCCCATGCTCCAGC 2700
2641 CCACACAGTTCCTCATCAGCACTGTCTCATGAGAGTGTGGGTCCCATGCTCCAGC 2700
2701 TGCCTCTTTGGCTCATGAGTTTTTCTGTTCTGCTGCCCCCACCACATGTCATTCCT 2760
2701 TGCCTCTTTGGCTCATGAGTTTTTCTGTTCTGCTGCCCCCACCACATGTCATTCCT 2760
2761 CAGATTTGCTGATTTCTATTTCTGTCGACCTCCCTGCTGCTTGGGATTTCTACTT 2820
2761 CAGATTTGCTGATTTCTATTTCTGTCGACCTCCCTGCTGCTTGGGATTTCTACTT 2820
2821 CTTCCTGTGGGGCCCATAGCTGTCTTAAAGAGTGAAGAAATGAATGAATGAATGA 2880
2821 CTTCCTGTGGGGCCCATAGCTGTCTTAAAGAGTGAAGAAATGAATGAATGAATGA 2880
2881 CTGGGGCCAGAAATCCATAAATGAGTGAAGAGTGTCTGCTGCTGCTTCTACCC 2940
2881 CTGGGGCCAGAAATCCATAAATGAGTGAAGAGTGTCTGCTGCTGCTTCTACCC 2940
2941 CCACTCCAGTACATAACTACTATGCTGTAGAGCCATTTCTATATGCTGAATGTTCTG 3000
2941 CCACTCCAGTACATAACTACTATGCTGTAGAGCCATTTCTATATGCTGAATGTTCTG 3000
3001 CTGTTGCAAACTTGCAGGGTATTAGCCAGTGTGTGCCAAGAGTGTTCGGGGGCAAC 3060
3001 CTGTTGCAAACTTGCAGGGTATTAGCCAGTGTGTGCCAAGAGTGTTCGGGGGCAAC 3060
3061 AGAATGACTCAGACCAAGATGAGTGTGTTAGGCTTTGCTTCTGCTGTTTTTCTT 3120
3061 AGAATGACTCAGACCAAGATGAGTGTGTTAGGCTTTGCTTCTGCTGTTTTTCTT 3120
3121 TGAATAGTCAATTTCTGAGGCTCCCTCATCTTCCATACCTAGGCCACTTTTAGCC 3180
3121 TGAATAGTCAATTTCTGAGGCTCCCTCATCTTCCATACCTAGGCCACTTTTAGCC 3180
3181 CTTACCTTAATCTCTCAGATGAGTGTTCACAAAGATGTTAAGTACTGAATCATGTG 3240
3181 CTTACCTTAATCTCTCAGATGAGTGTTCACAAAGATGTTAAGTACTGAATCATGTG 3240
3241 TGACTGAGACAGAGATGGCAATGAATGGCACCACCTTTCTCTCTCTCTCTCTCTCTCT 3300
3241 TGACTGAGACAGAGATGGCAATGAATGGCACCACCTTTCTCTCTCTCTCTCTCTCTCT 3300

Db 1921 GCAGAGTGAATTGAGCAATTTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCGCTTT 1980
QY 1981 CAGAGCTCAAAACAAGTAATACACCAAGGTCCTGACTGCAATTTGCTTGTGAGCAGGCT 2040
Db 1981 CAGAGCTCAAAACAAGTAATACACCAAGGTCCTGACTGCAATTTGCTTGTGAGCAGGCT 2040
QY 2041 TGTCTGTGCTCAGCTCAGGCTCCCTAGCTGCTTGGAGGCTCCCTTTGATTTCTAGACCTGG 2100
Db 2041 TGTCTGTGCTCAGCTCAGGCTCCCTAGCTGCTTGGAGGCTCCCTTTGATTTCTAGACCTGG 2100
QY 2101 AAAAGGTGCTCCTAGGACAGGCTCCCTGAGGCGCTCAGAGCTGGGATTTCTGCGCTGGA 2160
Db 2101 AAAAGGTGCTCCTAGGACAGGCTCCCTGAGGCGCTCAGAGCTGGGATTTCTGCGCTGGA 2160
QY 2161 ACAAGGACCTGAGGAATGTTTTGCTGGGATGATGTGCTGGTCAAGGACCCCTTGGGC 2220
Db 2161 ACAAGGACCTGAGGAATGTTTTGCTGGGATGATGTGCTGGTCAAGGACCCCTTGGGC 2220
QY 2221 ATCGCTTCCCTGCTTGGTGTGAGGACAGGACAGGCAATGATGCTTCAAGTAGCCT 2280
Db 2221 ATCGCTTCCCTGCTTGGTGTGAGGACAGGACAGGCAATGATGCTTCAAGTAGCCT 2280
QY 2281 TATCATTCACAGGTGCTCTAGCTGCAACAATGATGACAGAGATCACCAAGGA 2340
Db 2281 TATCATTCACAGGTGCTCTAGCTGCAACAATGATGACAGAGATCACCAAGGA 2340
QY 2341 TTAATTCGAGGTGTTTTCTTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTT 2400
Db 2341 TTAATTCGAGGTGTTTTCTTTATTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTT 2400
QY 2401 TTTTTTGGCAGACAGGTGTTTTGATTTAGGACCTTCAAGGAAAGGATGCTGTACC 2460
Db 2401 TTTTTTGGCAGACAGGTGTTTTGATTTAGGACCTTCAAGGAAAGGATGCTGTACC 2460
QY 2461 AGTGGTGGCTGGGTGCTGCGCTCCAGTGTCCACACCTCTTCAACACCCCACTTGGCTCC 2520
Db 2461 AGTGGTGGCTGGGTGCTGCGCTCCAGTGTCCACACCTCTTCAACACCCCACTTGGCTCC 2520
QY 2521 TTTTGCATCTGATGTGAGGTTTTCTGTTTGGTGAGATCAGGTTGTTTGGTAAAGA 2580
Db 2521 TTTTGCATCTGATGTGAGGTTTTCTGTTTGGTGAGATCAGGTTGTTTGGTAAAGA 2580
QY 2581 AAGGAAAGGCTTCTGATGCTTTGCAACAAGCTTACCTGTGGGTTTCACTGCTGAGAG 2640
Db 2581 AAGGAAAGGCTTCTGATGCTTTGCAACAAGCTTACCTGTGGGTTTCACTGCTGAGAG 2640
QY 2641 CCACCAACAGTTCCCATCAGCACTGTCTCCATGACAGAGTTGCTGGGTCCCAAGTCCAGC 2700
Db 2641 CCACCAACAGTTCCCATCAGCACTGTCTCCATGACAGAGTTGCTGGGTCCCAAGTCCAGC 2700
QY 2701 TGCTCTTTGGCTTCATGGTTTTCTGCTTCTGCTCCGCCCAACCCCACTGCTGCAATCCT 2760
Db 2701 TGCTCTTTGGCTTCATGGTTTTCTGCTTCTGCTCCGCCCAACCCCACTGCTGCAATCCT 2760
QY 2761 CAAGAATTTGCTGATTTCTTTCTGGCACCTCCCTGCTGCTTGGGGATTTCTACTT 2820
Db 2761 CAAGAATTTGCTGATTTCTTTCTGGCACCTCCCTGCTGCTTGGGGATTTCTACTT 2820
QY 2821 CTTTCTGTGGGGCCCATAGCTGTTGCTTACAGGTAGAAATGAAATGAACTATTGA 2880
Db 2821 CTTTCTGTGGGGCCCATAGCTGTTGCTTACAGGTAGAAATGAAATGAACTATTGA 2880
QY 2881 CTGGGCCCCAGAAATCCATAAATGCTGACAGAGTTGTTTTCTGTGCTCTTCTACCC 2940
Db 2881 CTGGGCCCCAGAAATCCATAAATGCTGACAGAGTTGTTTTCTGTGCTCTTCTACCC 2940
QY 2941 CACTCCAGTACATACTATGATGCTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
Db 2941 CCACTCCAGTACATACTATGATGCTGTAGAGCCATTCTATATGCTGAATGTTCTG 3000
QY 3001 CTGTTCCAACTTGCAGGGATTTAGCCAGTTGTTGTGCCAAGCAGTTTTCGGGACAAAC 3060

Db 3001 CTGTTCCAACTTGCAGGGATTTAGCCAGTTGTTGTGCCAAGCAGTTTTCGGGACAAAC 3060
QY 3061 AGAATGACTCAGACCAAGATGATAGGATGTTAGGCTTTGCTTCTTGTGTTTTCTT 3120
Db 3061 AGAATGACTCAGACCAAGATGATAGGATGTTAGGCTTTGCTTCTTGTGTTTTCTT 3120
QY 3121 TGAATAGTCAATTTGCTGCGAGGTCCTTCACTTCCATACCTAGCCCACTTTTAGCC 3180
Db 3121 TGAATAGTCAATTTGCTGCGAGTCCCTTCACTTCCATACCTAGCCCACTTTTAGCC 3180
QY 3181 CTTACCTTAAATCTCTCAGATAAGTTGGTTCAAAAGATGTTAAGTACTGAATCATGTG 3240
Db 3181 CTTACCTTAAATCTCTCAGATAAGTTGGTTCAAAAGATGTTAAGTACTGAATCATGTG 3240
QY 3241 TGACTGACACAGAGATGGCAATGAATGGACACCACTTTCTCCCTTCTGCGCCCAAGG 3300
Db 3241 TGACTGACACAGAGATGGCAATGAATGGACACCACTTTCTCCTTCTGCGCCCAAGG 3300
QY 3301 CAGGTACACATGATCTGATCAGAGTTGCTGCTATTTCTCTGCTGATCTTCACTACTA 3360
Db 3301 CAGGTACACATGATCTGATCAGAGTTGCTGCTATTTCTCTGCTGATCTTCACTACTA 3360
QY 3361 GGTGCGCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGC 3420
Db 3361 GGTGCGCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGC 3420
QY 3421 TGTCTGTGAAAGTGTTCCTCANTGACCTTTTCTTCCCTTTGAAATCCCTCTTGTCTGGAG 3480
Db 3421 TGTCTGTGAAAGTGTTCCTCANTGACCTTTTCTTCCCTTTGAAATCCCTCTTGTCTGGAG 3480
QY 3481 TAGTCTTGTGCTTCTTCTGCTCCAGTAGGCTTTTCTTACCCAGCCCTTGTGCGAGG 3540
Db 3481 TAGTCTTGTGCTTCTTCTGCTCCAGTAGGCTTTTCTTACCCAGCCCTTGTGCGAGG 3540
QY 3541 CTAAGCTGGTACAAAGAGCTGCCAATCACAAGATTTTGTAGGCGAGAGAGTGCAGGGA 3600
Db 3541 CTAAGCTGGTACAAAGAGCTGCCAATCACAAGATTTTGTAGGCGAGAGAGTGCAGGGA 3600
QY 3601 AGAGGACAGATATGACCTTCCCTTTGAAGAGAGGGAAGGCTTACAGTGGCCCA 3660
Db 3601 AGAGGACAGATATGACCTTCCCTTTGAAGAGAGGGAAGGCTTACAGTGGCCCA 3660
QY 3661 TAAATGCTGACTCACAACCTTCACTTCTTAAATGCTGTGAGGAGCTTGGAGCTGCTG 3720
Db 3661 TAAATGCTGACTCACAACCTTCACTTCTTAAATGCTGTGAGGAGCTTGGAGCTGCTG 3720
QY 3721 GATCCCAAGTGTGTGTAGGAGGCCACAGTACAGAGTGGGCCCAAGTGGGTTCCTCA 3780
Db 3721 GATCCCAAGTGTGTGTAGGAGGCCACAGTACAGAGTGGGCCCAAGTGGGTTCCTCA 3780
QY 3781 GGTCAAGAAATGTGGGCCCCCAGGCAAGGTGACGCTTTTGTCTCAGCTTCCATGATGCTA 3840
Db 3781 GGTCAAGAAATGTGGGCCCCCAGGCAAGGTGACGCTTTTGTCTCAGCTTCCATGATGCTA 3840
QY 3841 GACCTTCAGGCGAGTCTGACAGATGAGTTCCCTACCTTTTCTTCTTCTTCAATGACAAA 3900
Db 3841 GACCTTCAGGCGAGTCTGACAGATGAGTTCCCTACCTTTTCTTCTTCTTCAATGACAAA 3900
QY 3901 TCACCAATCACTACAGCTGCTGCTTCTTCTTCCAAAGTAGCCCAAGTCTCTGGGCA 3960
Db 3901 TCACCAATCACTACAGCTGCTGCTTCTTCTTCCAAAGTAGCCCAAGTCTCTGGGCA 3960
QY 3961 GATCAGGAGGAGTGCCTATCCATGAGTGAAGGCCAGTGTCTTCCCTCACCTGGGTGCTC 4020
Db 3961 GATCAGGAGGAGTGCCTATCCATGAGTGAAGGCCAGTGTCTTCCCTCACCTGGGTGCTC 4020
QY 4021 CACACTTGTGACCTCAGTTTTAGGACCCCAAGATCTGTGTGTTTTCTTAGATGCTAGC 4080
Db 4021 CACACTTGTGACCTCAGTTTTAGGACCCCAAGATCTGTGTGTTTTCTTAGATGCTAGC 4080
QY 4081 TTTTCTCCAGGGGACCAACAGCAGGTGAAGCTCAAGAGCGCATGGCTCTGCTAATAGTAA 4140
Db 4081 TTTTCTCCAGGGGACCAACAGCAGGTGAAGCTCAAGAGCGCATGGCTCTGCTAATAGTAA 4140

QY 4141 ATTGTTTTTCAGGGCCTTGTTCAGCTGAGAGCTTCATGTCACACAGATTCTGAGAGTTC 4200
Db 4141 ATTGTTTTTCAGGGCCTTGTTCAGCTGAGAGCTTCATGTCACACAGATTCTGAGAGTTC 4200
QY 4201 ACAGACATTTTTTTTTTTTATTTTGTGTTTGTGTTTTCATGAGTTATCGACCATGGCT 4260
Db 4201 ACAGACATTTTTTTTTTTTATTTTGTGTTTGTGTTTTCATGAGTTATCGACCATGGCT 4260
QY 4261 GAGCTCAGGACATTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTATTTAAACCCCTCC 4320
Db 4261 GAGCTCAGGACATTTCTGTAGGAGACTGTTATTTCTGTAAAGATGTTATTTAAACCCCTCC 4320
QY 4321 TCACGCCCATACGGTGGCCCTGAGGCTGACCCGAGGCGAGTGGAGCTGCTGGTGTTC 4380
Db 4321 TCACGCCCATACGGTGGCCCTGAGGCTGACCCGAGGCGAGTGGAGCTGCTGGTGTTC 4380
QY 4381 CACGGGGAGGCGCAAGGCTGCTGAGCTGATTCTTCCAGCTGCTGCCCGAGCTTTCGCG 4440
Db 4381 CACGGGGAGGCGCAAGGCTGCTGAGCTGATTCTTCCAGCTGCTGCCCGAGCTTTCGCG 4440
QY 4441 CTTGCACAGCAGAGGTTGTTCAACCCAGGAGCAGCCAGGACCTGCTCTTTCGCTTT 4500
Db 4441 CTTGCACAGCAGAGGTTGTTCAACCCAGGAGCAGCCAGGACCTGCTCTTTCGCTTT 4500
QY 4501 CTTGGGGAGGAGGCTGCTTCTGCTGCTGTAACCTGTTCTTATGSCCAGCAACCCGCG 4560
Db 4501 CTTGGGGAGGAGGCTGCTTCTGCTGCTGTAACCTGTTCTTATGSCCAGCAACCCGCG 4560
QY 4561 CACTCAGACTGTTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
Db 4561 CACTCAGACTGTTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
QY 4621 AGCCAGGACTGTTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
Db 4621 AGCCAGGACTGTTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
QY 4681 AAAAAAAAAAAAAA 4697
Db 4681 AAAAAAAAAAAAAA 4697

RESULT 3
ACC46153
ID ACC46153 standard; cDNA; 5085 BP.
XX ACC46153;
AC ACC46153;
DT 02-JUN-2003 (first entry)
XX Human dithp extracellular signalling protein-encoding cDNA.
XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW extracellular signalling; gene; ss.
XX Homo sapiens.
OS
XX WO200297031-A2.
PN
XX
PD 05-DEC-2002.
XX
XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefield Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
DR P-PSDB; ABR41210.
XX
PT Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS Claim 2; SEQ ID NO 74; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which has extracellular signalling activity. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5085 BP; 1178 A; 1310 C; 1407 G; 1190 T; 0 U; 0 Other;
Query Match 38.7%; Score 1819; DB 7; Length 5085;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3459; Conservative 0; Mismatches 10; Indels 11; Gaps 8;
QY 1 CAAAAGCAGCCAGGACAAACCGGAGCAGGACCCGAGGAGCCAGGAGCCGCGCCAGCC 60
Db 326 CAAAAGCAGCCAGGACAAACCGGAGCAGGACCCGAGGAGCCAGGAGCCGCGCCAGCC 385
QY 61 AGCGGGCTCTGCTAGTAGAAGAGGTCCCGCAGCAGCAGGCTCTCTGGAGCCGG 120
Db 386 AGCGGGCTCTGCTAGTAGAAGAGGTCCCGCAGCAGCAGGCTCTCTGGAGCCGG 445
QY 121 AGCGGGCTCAAGCAGGAGGCTGCTGCTGGGCGCTCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 446 AGCGGGCTCAAGCAGGAGGCTGCTGCTGGGCGCTCTGCTGCTGCTGCTGCTGCTGCTG 505
QY 181 GCCGCCAACTGGAGACATACCTGAGCACAATACCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

QY	2395	TTCTTTTTTTTTGGCACATGACAGTGTGTTGTATGATGAGGACCTTTCAAAGGAAAGGATGCG	2454
Db	2724	TTTTTTTTTTTTTGGCACATGACAGTGTGTTGTATGATGAGGACCTTTCAAAGGAAAGGATGCG	2783
QY	2455	TGTACAGTGGTGCCTGGGTGCGCTGGCTCCAGTGTCCACAGTCCACACCTCCCTTCACACCCACATTT	2514
Db	2784	TGTAGCAGTGGTGCCTGGGTGCGCTGGCTCCAGTGTCCACAGTCCACACCTCCCTTCACACCCACATTT	2843
QY	2515	GGCTCCTTTTGGCCATCTTGATGCTGAGGTTTTCTTGTTTGGTGAGATCAGGTGTGTTTGTGTT	2574
Db	2844	GGCTCCTTTTGGCCATCTTGATGCTGAGGTTTTCTTGTTTGGTGAGATCAGGTGTGTTTGTGTT	2903
QY	2575	AAAGAAGAAAGAGGCTTCTGATGCTTTGGCACAAAGCTTACTGTGGTGTTCAGTCTCT	2634
Db	2904	AAAGAAGAAAGAGGCTTCTGATGCTTTGGCACAAAGCTTACTGTGGTGTTCAGTCTCT	2963
QY	2635	GAGAGGCCACACAGTGTCCCATCAGCACTGTCTCCATGCAGCAGTGTGTGGGTGCCCATG	2694
Db	2964	GAGAGGCCACACAGTGTCCCATCAGCACTGTCTCCATGCAGCAGTGTGTGGGTGCCCATG	3023
QY	2695	TCAGTGTGCTCTTTGGCTTTCATGGGTTTTTGTCTTCGTGCCGCCACCCCCACATGTGCG	2754
Db	3024	TCAGTGTGCTCTTTGGCTTTCATGGGTTTTTGTCTTCGTGCCGCCACCCCCACATGTGCG	3083
QY	2755	AATCCTCAGATTTGCTCGATTCATTTCTCGCACCTCCCTCGCTGTCTCTTGGGGATT	2814
Db	3084	AATCCTCAGATTTGCTCGATTCATTTCTCGCACCTCCCTCGCTGTCTCTTGGGGATT	3143
QY	2815	CTACTTCTTCTCTGTGTGGG-GCCCATAGCTGTGTTCTTAACAGGTAAAGAAATGAATGA	2873
Db	3144	CTACTTCTTCTCTGTGTGGGAGCCATAGCTGTGTTCTTAACAGGTAAAGAAATGAATGA	3203
QY	2874	CTATTGACTGGGCCCCAGAAATCCATAAATGSGCTGCAGACAGTTGTTTCTGTGCTCTGT	2933
Db	3204	CTATTGACTGGGCCCCAGAAATCCATAAATGSGCTGCAGACAGTTGTTTCTGTGCTCTGT	3263
QY	2934	TCTACCCCCACTCCAGTACATAACTATATGTACTGTGTAGAGCCATTCATATGCTGAA	2993
Db	3264	TCTACCCCCACTCCAGTACATAACTATATGTACTGTGTAGAGCCATTCATATGCTGAA	3323
QY	2994	TGTTCTGCTGTTGAAACTTGGCCAGGGTATTAGCCAGTGTGTTGTGCCAAGCAGTTTTCGG	3053
Db	3324	TGTTCTGCTGTTGAAACTTGGCCAGGGTATTAGCCAGTGTGTTGTGCCAAGCAGTTTTCGTG	3383
QY	3054	GGACACAGAAATGACTCAGACCAAGATGATAGATGGTTTAGGGCTTTGCTTCTTGCTGT	3113
Db	3384	GGACACAGAAATGACTCAGACCAAGATGATAGATGGTTTAGGGCTTTGCTTCTTGCTGT	3443
QY	3114	TTTTTCTTTGAA-CTAG-TCATTGCTCGAGGTCCCTTCATCTTCCATACCTAGCCOACT	3171
Db	3444	TTTTTCTTTGAACTAGTTTCATTGCTCGAGGTCCCTTCATCTTCCATACCTAGCCOACT	3503
QY	3172	CTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCAAGAAGATCTTAAGTACTG	3231
Db	3504	CTTTTAGCCCTTACCTTAAATCTCTCAGATAAGTTGGTTCAAGAAGATCTTAAGTACTG	3563
QY	3232	AATCATGTGTGACTGAGACACAGATGGCAAAATGAATGGCACACCAATTTCTCCCTTCTCCT	3291
Db	3564	AATCATGTGTGACTGAGACACAGATGGCAAAATGAATGGCACACCAATTTCTCCCTTCTCCT	3623
QY	3292	GGCCAGGAGATGACACCTGATCAGATTCGCTTGCCTGCTATTTCTCTGGTGTATCCT	3351
Db	3624	GGCCAGGAGATGACACCTGATCAGATTCGCTTGCCTGCTATTTCTCTGGTGTATCCT	3683
QY	3352	TCACATCTAGTGTGCCCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTG	3411
Db	3684	TCACATCTAGTGTGCCCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTG	3743
QY	3412	AGATPACTGTCTCTGTGAAGTGTTCCTCATGACCTTTTCTTCTCCCTTTGAATCCCTCT	3471
Db	3744	AGATPACTGTCTCTGTGAAGTGTTCCTCATGACCTTTTCTTCTCCCTTTGAATCCCTCT	3803

RESULT 4
ACC46130 standard; cDNA; 6577 BP.
XX AC ACC46130;
XX DT 02-JUN-2003 (first entry)
XX DE Human dithp extracellular signalling protein-encoding cDNA.
XX KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW extracellular signalling; gene; ss.
XX XX
OS Homo sapiens.
XX XX
PN W02C00297031-A2.
XX XX
PD 05-DEC-2002.
XX XX
PF 27-MAR-2002; 2002WO-US010056.
XX XX
PR 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0239428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX XX
PA (INCY-) INCYTE GENOMICS INC.
XX PI Daifo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI DuFour GE, Hillman JL, Yu JY, Tusson O, Yap PE, Amshey SR;
PI Daugherty SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RV, Urashka ME;
XX XX
DR WPI; 2003-129518/12.
DR P-FSDB; ABR41186.
XX XX
PT Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX XX
PS Claim 2; SEQ ID NO 51; 591pp; English.
XX XX
CC The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP

CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a dithp cDNA encoding a DITHP protein
 CC which has extracellular signalling activity. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 6577 BP; 1496 A; 1701 C; 1836 G; 1540 T; 0 U; 4 Other;
 Query Match 38.7%; Score 1819; DB 7; Length 6577;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 3459; Conservative 0; Mismatches 10; Indels 11; Gaps 8;
 QY 1 CAAAAAGCCAGCCAGGACCAACCGGAAGCAGGACCCCGAGGAGCCAGGAGCCCGCCAGCC 60
 DB 326 CAAAAAGCCAGCCAGGACCAACCGGAAGCAGGACCCCGAGGAGCCAGGAGCCCGCCAGCC 385
 QY 61 AGCGGCTCTGCTAGTAGAGAGAGGTCCTCGGAGAGCCAGGTCCTCTCGAGAGCCGG 120
 DB 386 AGCGGCTCTGCTAGTAGAGAGAGGTCCTCGGAGAGCCAGGTCCTCTCGAGAGCCGG 445
 QY 121 AGCGGCTCTGCTAGTAGAGAGGTCCTCGGAGAGCCAGGTCCTCTGAGAGGTCGA 180
 DB 446 AGCGGCTCTGCTAGTAGAGAGGTCCTCGGAGAGCCAGGTCCTCTGAGAGGTCGA 505
 QY 181 GCGCCAACTGGAAGACATACCTGAGCACATATCTGTGTGACAAATACCGAGGAGCCCGG 240
 DB 506 GCGCCAACTGGAAGACATACCTGAGCACATATCTGTGTGACAAATACCGAGGAGCCCGG 565
 QY 241 GCGAGAGTGGGACAGGTCAGCGGCTGAACCGGAGATCGAGAGAGTCCCGGACCT 300
 DB 566 GCGAGAGTGGGACAGGTCAGCGGCTGAACCGGAGATCGAGAGAGTCCCGGACCT 625
 QY 301 ATGTGCAAGGAATGGGAGCCTGAACCAACTCCAGTCTATGGAGAGAGGAAACCT 360
 DB 626 ATGTGCAAGGAATGGGAGCCTGAACCAACTCCAGTCTATGGAGAGAGGAAACCT 685
 QY 361 CAAAGGGGNTCAACACAGAGAGATCCGACAGATCCGAGGTCGAGAGCCGAGAC 420
 DB 686 CAAAGGGGNTCAACACAGAGAGATCCGACAGATCCGAGGTCGAGAGCCGAGAC 745
 QY 421 ATCGAAGGCCACAGGAGAGAGAGAGAGGTTTGGGAGGAGATCACTGTTGCTGA 480
 DB 746 ATCGAAGGCCACAGGAGAGAGAGAGAGGTTTGGGAGGAGATCACTGTTGCTGA 805
 QY 481 TGACAGATTTGAATCTCTGATACCCAGAGAGAGTCTGCTCTGTCAGAGAGT 540
 DB 806 TGACAGATTTGAATCTCTGATACCCAGAGAGAGTCTGCTCTGTCAGAGAGT 865
 QY 541 ATGCTCAAATGCTGGAGGAGACCGGAATTCACAGAGCAGATGAAGCTCTTACAGAAAA 600
 DB 866 ATGCTCAAATGCTGGAGGAGACCGGAATTCACAGAGCAGATGAAGCTCTTACAGAAAA 925
 QY 601 AGCAGAGCAGTGTGTGAAGAGAGGACCACTGCGGCGTGTGAGCACAGAGCCGCTCC 660
 DB 926 AGCAGAGCAGTGTGTGAAGAGAGGACCACTGCGGCGTGTGAGCACAGAGCCGCTCC 985
 QY 661 TGGCCCGCAGCAGCTTGTGAGAGCTATGCGGTGAGCTGAGCGGCACACCGCTCCCTCA 720
 DB 986 TGGCCCGCAGCAGCTTGTGAGAGCTATGCGGTGAGCTGAGCGGCACACCGCTCCCTCA 1045
 QY 721 AGGAAGAGGTGTGACGCGGCGCGGAGGAGAGAGAGCGCAAGAGAGTGTGCTCGC 780
 DB 1046 AGGAAGAGGTGTGACGCGGCGCGGAGGAGAGAGAGCGCAAGAGAGTGTGCTCGC 1105
 QY 781 ACTTCCAGTGTACATGATGACATTCAGCTCAGATGGAACAGACCAATGAGCGCAACT 840
 DB 1105 ACTTCCAGTGTACATGATGACATTCAGCTCAGATGGAACAGACCAATGAGCGCAACT 840

DB 1106 ACTTCCAGTGTACATGATGACATTCAGCTCAGATGGAACAGACCAATGAGCGCAACT 1165
 QY 841 CCAAGCTCGCCAGAGAGAAATGAGCTGCTGAGAGCTCAAGAGCTCAATGAGCAGT 900
 DB 1166 CCAAGCTCGCCAGAGAGAAATGAGCTGCTGAGAGCTCAAGAGCTCAATGAGCAGT 1225
 QY 901 ATGAGCTCGCGAGGAGCATATCGCAAAAGTCTTCAAAACACAGGACCTCAACAGCAGC 960
 DB 1226 ATGAGCTCGCGAGGAGCATATCGCAAAAGTCTTCAAAACACAGGACCTCAACAGCAGC 1285
 QY 961 TGTGTGATGCTCAAGCTCCAGAGCCCGCAGGAGATGCTAAAGAGGAGCAAGAGCGGACCC 1020
 DB 1286 TGTGTGATGCTCAAGCTCCAGAGCCCGCAGGAGATGCTAAAGAGGAGCAAGAGCGGACCC 1345
 QY 1021 AGCGGAGAGAGGATTTCTCTCTGAAGAGGAGCTAGAGTCCAGAGGATGCTGTGAGCTGA 1080
 DB 1346 AGCGGAGAGAGGATTTCTCTCTGAAGAGGAGCTAGAGTCCAGAGGATGCTGTGAGCTGA 1405
 QY 1081 TGAAGCAGAGAGAGACCCACTCTGAAGCAACAGCTTGGCTTATACAGAGAACTTTGAGG 1140
 DB 1406 TGAAGCAGAGAGAGACCCACTCTGAAGCAACAGCTTGGCTTATACAGAGAACTTTGAGG 1465
 QY 1141 AGTTCCAGAGACACTTTCCAAAGAGCAGGAGTATTCACCACTTCAAGCAGAGAGATGG 1200
 DB 1466 AGTTCCAGAGACACTTTCCAAAGAGCAGGAGTATTCACCACTTCAAGCAGAGAGATGG 1525
 QY 1201 AAAAGATGACTTAAGAGATCAAGAGCTGGAGAGAGAAACCACTATGTATCCCGTCCCGGT 1260
 DB 1526 AAAAGATGACTTAAGAGATCAAGAGCTGGAGAGAGAAACCACTATGTATCCCGTCCCGGT 1585
 QY 1261 GGGAGAGCAGCAACAGGCCCTCTGAGATGCTGAGGAGAGAAACAGTCCCGGATAAG 1320
 DB 1586 GGGAGAGCAGCAACAGGCCCTCTGAGATGCTGAGGAGAGAAACAGTCCCGGATAAG 1645
 QY 1321 AACTGGAGGGCTTGCAGGTAAAAATCCAAGCTGGAGAGAGCTGTGCGCGGCACTCAGA 1380
 DB 1646 AACTGGAGGGCTTGCAGGTAAAAATCCAAGCTGGAGAGAGCTGTGCGCGGCACTCAGA 1705
 QY 1381 CAGAGCCGAATGACTTGAACAGAGGGTACAGACCTGAGTGTGCTGTGGCAGGGCTCCC 1440
 DB 1706 CAGAGCCGAATGACTTGAACAGAGGGTACAGACCTGAGTGTGCTGTGGCAGGGCTCCC 1765
 QY 1441 TCACTGACAGTGGCCCTGAGAGAGGAGCCAGAGGGCTTGGGGCTCAAGCAGCCAGTCCC 1500
 DB 1766 TCACTGACAGTGGCCCTGAGAGAGGAGCCAGAGGGCTTGGGGCTCAAGCAGCCAGTCCC 1825
 QY 1501 CCAAGGTTCACAGAGCGCCTTGTATCCAGGAGCAACGAGCAGAGAGATCAGGCGAGA 1560
 DB 1826 CCAAGGTTCACAGAGCGCCTTGTATCCAGGAGCAACGAGCAGAGAGATCAGGCGAGA 1885
 QY 1561 CTGGGCTCAAGAGCCACCTCCGAGGGCTTGAAGAGCTGTGGTGTGGTCACTGCTG 1620
 DB 1886 CTGGGCTCAAGAGCCACCTCCGAGGGCTTGAAGAGCTGTGGTGTGGTCACTGCTG 1945
 QY 1621 GAAGGAGCGGCGAGCCCGCAGGCGCTTGGCCCAATAAAGCTCCCATGTGTGAGCAGCCCA 1680
 DB 1946 GAAGGAGCGGCGAGCCCGCAGGCGCTTGGCCCAATAAAGCTCCCATGTGTGAGCAGCCCA 2005
 QY 1681 TTGCTGAAGCAGAGATGTTCTTGAACCTGGCTGAGCTGTGGCACTTGCATTTTGAATTTT 1740
 DB 2006 TTGCTGAAGCAGAGATGTTCTTGAACCTGGCTGAGCTGTGGCACTTGCATTTTGAATTTT 2064
 QY 1741 GTGGGTCAAGTACATAGAGGAGTGTGCAAGGCTTGAAGGCTTGAAGTATGATTTATACCTG 1800
 DB 2065 GTGGGTCAAGTACATAGAGGAGTGTGCAAGGCTTGAAGGCTTGAAGTATGATTTATACCTG 2124
 QY 1801 TAAGTGTACAGTGGGCTTGCATTTGGGGATGGGGGTGTGACAGATGAAGTCAAGTGTGCTG 1860
 DB 2125 TAAGTGTACAGTGGGCTTGCATTTGGGGATGGGGGTGTGACAGATGAAGTCAAGTGTGCTG 2184
 QY 1861 TCTGTGAGCTGAAGAGTCTTTGAGAGGGGCTGTCTATCTGTAGCTGCCATCAGAGTGTGTTG 1920
 DB 2185 TCTGTGAGCTGAAGAGTCTTTGAGAGGGGCTGTCTATCTGTAGCTGCCATCAGAGTGTGTTG 2244

QY	1921	GCAGAGTGACTTGAGCATTTCTCTGTCTGATTTTGAGGCTCAGACCCCTCCCTGCCCCTTT	1981
DB	2245	GCAAGAGTGACTTGAGCATTTCTCTGTCTGATTTGAGGCTCAGACCCCTCCCTGCCCC- TT	2303
QY	1981	CAGAGCTCAAAACAAGTAATAACACCAAGGCTCTGACTGCATTTGTCTTGTGAGCAGGGCT	2040
DB	2304	CAGAGCTCAAGACAAGTAATAACACCCAGGCTCTGACTGCATTTGTCTTGTGAGCAGGGCT	2363
QY	2041	TGCTTGTGCTCAGCTCAGGCCCCCTCTAGCTGCT- TGAGGCTCTCTTTGATTTCTTAGACCTG	2099
DB	2364	TGCTTGTGCTCAGCTCAGGCCCCCTCTAGCTGCTCTGAGGCTCTTTGATTTCTCTAGACCTG	2423
QY	2100	GAAGAAGTGTCCCTAGGCAGAGCCCTGGCAGGGGGCTCAGAGCT- GGGATTTCTCTGCTG	2158
DB	2424	GAAGAAGTGTCCCTAGGCAGAGCCCTGGCAGGGGGCTCAGAGCTGGGGATTTCTGCTG	2483
QY	2159	GAACAAGGAGCTGGAGAAATGTTTTTGGGTGGGATGATGTCTGTCAGGAGCCCCCTGG	2218
DB	2484	GAACAAGGAGCTGGAGAAATGTTTTTGGGTGGGATGATGTCTGTCAGGAGCCCCCTGG	2543
QY	2219	GCATCGCTTCCCTCGCCCTTGGTGTAGTGCAGGACCAAGGCCAATGATGCTTCTCAGTAGC	2278
DB	2544	GCATCGCTTCCCTCGCCCTTGGTGTAGTGCAGGACCAAGGCCAATGATGCTTCTCAGTAGC	2603
QY	2279	CTTATCANTCAGAGTGCCCTCTCTAGCCTGCACAAATGATGACAAGAGATCACCCAAAG	2338
DB	2604	CTTATCANTCAGAGTGCCCTCTCTAGCCTGCACAAATGATGACAAGAGATCACCCAAAG	2663
QY	2339	GATTAATTCGAAAGTGTTTTTTTCTTTTATTTCTTTTCTTTTCTTTTCTTTTCTTTT	2394
DB	2664	GATTAATTCGAAAGTGTTTTTTTCTTTTAAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	2723
QY	2395	TTCTTTTTTTTTTGCAATCAGAGTGTTTGTATGAGACCTTCCAGGAAAAAGGATGC	2454
DB	2724	TTTTTTTTTTTTTTTGCAATCAGAGTGTTTGTATGAGACCTTCCAGGAAAAAGGATGC	2783
QY	2455	TGTACAGTGGTGCCTGGGTCCCTGGCTCCAGTGCACAGCTTACTGTGGGTTTCAGTCCCT	2514
DB	2784	TGTACAGTGGTGCCTGGGTCCCTGGCTCCAGTGCACAGCTTACTGTGGGTTTCAGTCCCT	2843
QY	2515	GGCTCCTTTGCCATCTTGATGCTGAGGTTTTCTGTGTTGGTGAGATCAGGTTGTTGTGGT	2574
DB	2844	GGCTCCTTTGCCATCTTGATGCTGAGGTTTTCTGTGTTGGTGAGATCAGGTTGTTGTGGT	2903
QY	2575	AAAGAAGAAAGAGGCTTCTGATGGCTTTGCCACAAGCTTACTGTGGGTTTCAGTCCCT	2634
DB	2904	AAAGAAGAAAGAGGCTTCTGATGGCTTTGCCACAAGCTTACTGTGGGTTTCAGTCCCT	2963
QY	2635	GAGAGGCCACACACAGTTCCCATCAGCACTGTCTCCATGCAGCAGTTGCTGGGTGCCCATG	2694
DB	2964	GAGAGGCCACACACAGTTCCCATCAGCACTGTCTCCATGCAGCAGTTGCTGGGTGCCCATG	3023
QY	2695	TCCAGTGCCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCTGCCCCACACCCACATGTGC	2754
DB	3024	TCCAGTGCCTCTTTGGCTTCATGGGTTTTTCTGCTTCTGCTGCCCCACACCCACATGTGC	3083
QY	2755	AATCCTCAAGATTTGCTGATTTCTATTTCTGGCACCTCCCTCCCTGCTCTTTGGGGATT	2814
DB	3084	AATCCTCAAGATTTGCTGATTTCTATTTCCCTGGCACCTCCCTGCTCTTTGGGGATT	3143
QY	2815	CTACTTCTTCTGTGTGGG- GCCCATAGCTGTTGTCTTAACAGGTAAAGAAATGAA	2873
DB	3144	CTACTTCTTCTGTGTGGGAGCCCATAGTGTGTCTTAACAGGTAAAGAAATGAA	3203
QY	2874	CTATTGACTGGGCCCCCAAGAAATCCATAAATGGCTGCAGACAGTTGTTTCTGTGCTCTGT	2933
DB	3204	CTATTGACTGGGCCCCCAAGAAATCCATAAATGGCTGCAGACAGTTGTTTCTGTGCTCTGT	3263
QY	2934	TCTACCCCACTCCAGTACATAACTACTATGTAATCTGTGTAGGCCAATCTATATGCTGAA	2993
DB	3264	TCTACCCCACTCCAGTACATAACTACTATGTAATCTGTGTAGGCCAATCTATATGCTGAA	3323

Qy	2994	TGTTCTGCTGTTGCAAACTTGCACGGGTATTAGCCAGTGTTTGTGCAACGAGTTTTCGG	3053
Db	3324	TGTTCTGCTGTTGCAAACTTGCACGGGTATTAGCCAGTGTTTGTGCAACGAGTTTTCGT	3383
Qy	3054	GGCAACAGAGTAATCACTCAGACCAAGATGGATAGGATGGTTAGGGCTTTTGCTTCTTGCTGT	3113
Db	3384	GGCAACAGAGTAATCACTCAGACCAAGATGGATAGGATGGTTAGGGCTTTTGCTTCTTGCTGT	3443
Qy	3114	TTTTCTTTTGAATCTAG-TCATTTGCTGTCAGGTCCTTTCATCTTCCATCACTAGCCCACT	3171
Db	3444	TTTTCTTTTGAAGCTAGTTCATTTGCTGTCAGGTCCTTTCATCTTCCATCACTAGCCCACT	3503
Qy	3172	CTTTTAGCCCTTACCTTTAAATCTCTCAGATAAGTTGGTTTCCAAAGAAGTCTTAAGTACTG	3231
Db	3504	CTTTTAGCCCTTACCTTTAAATCTCTCAGATAAGTTGGTTTCCAAAGAAGTCTTAAGTACTG	3563
Qy	3232	AATCATGTGTGACTGTAGAACAGAGATGGCAAAATGAATGGCACCACTTTCTCCTTCTCCT	3291
Db	3564	AATCATGTGTGACTGTAGAACAGAGATGGCAAAATGAATGGCACCACTTTCTCCTTCTCCT	3623
Qy	3292	GCCCCAGGGCAGGTACCACTGATCTGCATCAGAGTTGCCGTGCTATCTCTGCTGTTATCCT	3351
Db	3624	GCCCCAGGGCAGGTACCACTGATCTGCATCAGAGTTGCCGTGCTATCTCTGCTGTTATCCT	3683
Qy	3352	TCACATCTTAGTGCCTTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTGCCATCTCTGGCTG	3411
Db	3684	TCACATCTTAGTGCCTTCAAGCAGCTGTGTGAGTGTGTGAGATCTCTGCCATCTCTGGCTG	3743
Qy	3412	AGATACTGCTGCTGCTGTGAAGTCTTTCCCATGACCTTTTCTTCCCTTTTGAATCCCTCT	3471
Db	3744	AGATACTGCTGCTGCTGTGAAGTCTTTCCCATGACCTTTTCTTCCCTTTTGAATCCCTCT	3803
RESULT 5			
ABA09008			
ID	ABA09008 standard; cDNA; 2523 BP.		
XX			
AC	ABA09008;		
XX			
DT	11-JAN-2002 (first entry)		
XX			
XX	Human LDL binding protein homologue-encoding cDNA, SEQ ID NO:784.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; growth factor;		
KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;		
KW	inhibin; chemokinesis; chemokinesis; thrombolysis; oncogenesis;		
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;		
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;		
KW	chronic inflammatory condition; proliferative retinopathy;		
KW	atherosclerosis; coronary heart disease; arterial ischaemia;		
KW	bone disorder; osteoporosis; vascular growth disorder;		
KW	tissue regeneration; wound healing; infection; immune disorder;		
KW	cell culture; drug screening; gene therapy; anti-inflammatory;		
KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;		
KW	cytostatic; osteoparathic; vasotropic; cardiant; virucide; antibacterial;		
KW	antifungal; vulnery; antitumor; ss.		

XX	CS	Homio sapiens.
XX	PN	WO200157188-A2.
XX	XX	
XX	PD	09-AUG-2001.
XX	XX	
XX	PF	05-FEB-2001; 2001WO-US003800.
XX	XX	
XX	PR	03-FEB-2000; 2000US-00496914.
XX	PR	27-APR-2000; 2000US-00560875.
XX	PA	(HYSE-) HYSEQ INC.
XX	XX	
PI	PI	Tang YT, Liu C, Drmanac RT;
XX	XX	
DR	DR	WPI; 2001-457740/49.

DR XX

PT

Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer.

PT

XX

PS

XX

CC

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABB08225-ABB09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders.

Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention

Query Match 33.8%; Score 1589; DB 4; Length 2523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAAAGCAGCCAGGACAAACCGAAGCAGGAGCCAGGAGCCAGGAGCGGCCAGCC 60
DB 165 CAAAAGCAGCCAGGACAAACCGAAGCAGGAGCCAGGAGCCAGGAGCGGCCAGCC 224

QY 61 AGCGGCTCTCAGTAGAGAGAGAGGTCCTCGGAGCAGCAGGCTCTCGAAGCCGG 120
DB 225 AGCGGCTCTCAGTAGAGAGAGAGGTCCTCGGAGCAGCAGGCTCTCGAAGCCGG 284

QY 121 AGGGGGCTCAAGCCAGAACCGGCTCAGTCTGGGGCCCTCTGTGATGTCTCTGAGGAGCTGA 180
DB 285 AGGGGGCTCAAGCCAGAACCGGCTCAGTCTGGGGCCCTCTGTGATGTCTCTGAGGAGCTGA 344

QY 181 GCGGCCAATCTGGAAGACATCTAGACACATCTGTGTGGACATATACCAAGGGGCCCCG 240
DB 345 GCGGCCAATCTGGAAGACATCTAGACACATCTGTGTGGACATATACCAAGGGGCCCCG 404

QY 241 GCGAGGATGGGCAACGGGTGAGCCGGCTGAAACCGAAGATGACAGAAATGCCCGGACCT 300
DB 405 GCGAGGATGGGCAACGGGTGAGCCGGCTGAAACCGAAGATGACAGAAATGCCCGGACCT 464

QY 301 ATGTGGCAAGGATGGGGAGCCCTGAACCAACTCCAGTAGTCTATGAGAGAGAACCCCT 360

DB 465 ATGTGGCAAGGATGGGGAGCCCTGAACCAACTCCAGTAGTCTATGAGAGAGAACCCCT 524

QY 361 CCAAGGGGATCCAAACACAGAAAGAGATCCGGCAGAGTACCGAGGTCCGAGACCCGAGACC 420

DB 525 CCAAGGGGATCCAAACACAGAAAGAGATCCGGCAGAGTACCGAGGTCCGAGACCCGAGACC 584

QY 421 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGGTTTGGGGAGGAGATCAGTTGCTGA 480

DB 585 ATCGAAGGCCACAGGAGAAAGAAAGCAAGGGTTTGGGGAGGAGATCAGTTGCTGA 644

QY 481 TGCAGACATTGATATCTCTGAGTACCCAGAGAGAGTGGCTCTGTGTCAAGAACT 540

DB 645 TGCAGACATTGATATCTCTGAGTACCCAGAGAGAGTGGCTCTGTGTCAAGAACT 704

QY 541 ATGTGAACTGTCTGGAGGAGCACCGGAATTTCAGAAAGCAGATGAAGTCTCTACAGAAA 600

DB 705 ATGTGAACTGTCTGGAGGAGCACCGGAATTTCAGAAAGCAGATGAAGTCTCTACAGAAA 764

QY 601 AGCAGACAGCTGTGTCAGAGAGAGAGACCACTGCGCGGTGAGCAGACAGAGCGCTCC 660

DB 765 AGCAGACAGCTGTGTCAGAGAGAGAGACCACTGCGCGGTGAGCAGACAGAGCGCTCC 824

QY 661 TGGCCCGCAGCAAGCTGTAGAGGCTATGCGGTGAGCTGCAGCGGCAACACCGCTCCCTCA 720

DB 825 TGGCCCGCAGCAAGCTGTAGAGGCTATGCGGTGAGCTGCAGCGGCAACACCGCTCCCTCA 884

QY 721 AGGAAGAGTGTGTCAGAGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

DB 885 AGGAAGAGTGTGTCAGAGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944

QY 781 ACTTCCAGGTGACACTGAATGACATTGACAGTGCAGATGGAACAGACAACTAGCGCACT 840

DB 945 ACTTCCAGGTGACACTGAATGACATTGACAGTGCAGATGGAACAGACAACTAGCGCACT 1004

QY 841 CCAAGCTCGCCCAAGAGAAACATGGAGTGTGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 900

DB 1005 CCAAGCTCGCCCAAGAGAAACATGGAGTGTGCTGAGAGGCTCAAGAGCTGATTGAGCAGT 1064

QY 901 ATGAGCTCGCGGAGGAGCATATCGACAAGTCTTCAACACAAAGGACCTACACAGCAGC 960

DB 1065 ATGAGCTCGCGGAGGAGCATATCGACAAGTCTTCAACACAAAGGACCTACACAGCAGC 1124

QY 961 TGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTTAAAGAGGAGGAGAGAGAGCGGACC 1020

DB 1125 TGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTTAAAGAGGAGGAGAGAGAGCGGACC 1184

QY 1021 AGCGGAGAGAGGATTTCTCTGAAAGAGGAGGAGTGTAGTCCAGAGGATGTTGAGCTGA 1080

DB 1185 AGCGGAGAGAGGATTTCTCTGAAAGAGGAGGAGTGTAGTCCAGAGGATGTTGAGCTGA 1244

QY 1081 TGAAGCAGCAGAGAGACCCACTGAAAGCAACAGCTTGGCCCTATACACAGAGAGTTGAGG 1140

DB 1245 TGAAGCAGCAGAGAGACCCACTGAAAGCAACAGCTTGGCCCTATACACAGAGAGTTGAGG 1304

QY 1141 AGTTCCAGAAACACTTTCCAAAAGCAGCGAGGTATTTCACCACATTCACAGCAGAGATGG 1200

DB 1305 AGTTCCAGAAACACTTTCCAAAAGCAGCGAGGTATTTCACCACATTCACAGCAGAGATGG 1364

QY 1201 AAAAGATGACTTAAGAAATCAAGAAAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGGT 1260

DB 1365 AAAAGATGACTTAAGAAATCAAGAAAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGGT 1424

QY 1261 GGGAGAGCAGCAACAGGCCCTGTGTGAGTGGCTGAGGAGAAACAGTCCGGATTAAG 1320

DB 1425 GGGAGAGCAGCAACAGGCCCTGTGTGAGTGGCTGAGGAGAAACAGTCCGGATTAAG 1484

QY 1321 AACTGGAGGCTCGAGTAAATCCAAAGCTGGAGAGAGCTGTGCGGGCACTGCGAGA 1380

DB 1485 AACTGGAGGCTCGAGTAAATCCAAAGCTGGAGAGAGCTGTGCGGGCACTGCGAGA 1544

QY 1381 CAGAGCGCAATGACCTGAACAAAGAGGGGTACAGGACCTGAGTGTGGTGGCGCAGGGCTCCC 1440

Db 1545 CAGAGCGCAATGACCTTGAACAAGAGGCTACAGGACCTGAGTGTGGTGGCCAGGCTCCC 1604
Qy 1441 TCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGCAACCCAGTCCC 1500
Db 1505 TCACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGCAACCCAGTCCC 1664
Qy 1501 CCAGGGTTCACAGAAGCGCCCTTGTCTACCCAGGAGCACCAGACAGAGCATCAGGCCAGA 1560
Db 1665 CCAGGGTTCACAGAAGCGCCCTTGTCTACCCAGGAGCACCAGACAGAGCATCAGGCCAGA 1724
Qy 1561 CTGGGCTCAAGAGCCGACCTCCGACAGGCGCTTAGAGAGCCTGGTGTGGTCACTGG 1620
Db 1725 CTGGGCTCAAGAGCCGACCTCCGACAGGCGCTTAGAGAGCCTGGTGTGGTCACTGG 1784
Qy 1621 GAAGGAGCGGCGAGCCAGC 1640
Db 1785 GAAGGAGCGGCGAGCCAGC 1804

RESULT 6
AAK52874
ID AAK52874 standard; cDNA; 2523 BP.
XX AC AAK52874;
XX DT 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 2403.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US004098.
XX PR 03-FEB-2000; 2000US-00496914.
XX PR 27-APR-2000; 2000US-00560875.
XX PR 20-JUN-2000; 2000US-00598075.
XX PR 19-JUL-2000; 2000US-00620325.
XX PR 01-SEP-2000; 2000US-00654936.
XX PR 15-SEP-2000; 2000US-00663561.
XX PR 20-OCT-2000; 2000US-00693325.
XX PR 30-NOV-2000; 2000US-00728422.
XX PA (HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
Xue AJ, Yang Y, Wejhrman T, Goodrich R;
WPI; 2001-476283/51.
P-FSDB; AAM79741.
Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.
Claim 1; Page 4687-4688; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 2523 BP; 666 A; 626 C; 794 G; 437 T; 0 U; 0 Other;
Query Match 33.8%; Score 1589; DB 4; Length 2523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAAAAGCAGCCAGGACCAACCGGAAGCAGAGCACCAGAGGAGCCAGAGGCGGCCAGCC 60
Db 165 CAAAAGCAGCCAGGACCAACCGGAAGCAGAGCACCAGAGGAGCCAGAGGCGGCCAGCC 224
Qy 61 AGGCGGCTCTGCAGTAGAAGCAGAGGTCCCGCAGCAGCCAGGCTCTCGGAGCGG 120
Db 225 AGGCGGCTCTGCAGTAGAAGCAGAGGTCCCGCAGCAGCCAGGCTCTCGGAGCGG 284
Qy 121 AGGCGGCTCAAGCCAGAAACCGCTCAGTCTGGGGCCCTTCGTGTGTCTCTGAGAGCTGA 180
Db 285 AGGCGGCTCAAGCCAGAAACCGCTCAGTCTGGGGCCCTTCGTGTGTCTCTGAGAGCTGA 344
Qy 181 GCCGCCAACTGGAAGACATCTGAGCAGCATCTGTGTGGACATTAACACAGGGGGGCCCG 240
Db 345 GCCGCCAACTGGAAGACATCTGAGCAGCATCTGTGTGGACATTAACACAGGGGGGCCCG 404
Qy 241 GCGAGGATGGGGCAGAGGCTGAGCCGCTGAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 300
Db 405 GCGAGGATGGGGCAGAGGCTGAGCCGCTGAACCCGGAAGATGCAGAGAAAGTCCCGGACCT 464
Qy 301 ATGTGGCAAGGATGGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGGAAACCT 360
Db 465 ATGTGGCAAGGATGGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGGAAACCT 524
Qy 361 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGTGCAGAGTGCAGAGCCGAGAGC 420
Db 525 CCAAGGGGATCCAAACACAGAGAGATCCGCGAGTGCAGAGTGCAGAGCCGAGAGC 584
Qy 421 ATCGAGGCCACAGGAGAGAGAAAGGTTGGGGAGAGGATCACCTGTGCTGA 480
Db 585 ATCGAGGCCACAGGAGAGAGAAAGGTTGGGGAGAGGATCACCTGTGCTGA 644
Qy 481 TGCAGCATTGAATCTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGT 540
Db 645 TGCAGCATTGAATCTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGT 704
Qy 541 ATGCTGAATCTGAGAGAGCACCAGGATTCACAGAGCAGATGAAGCTCCTACAGAAA 600
Db 705 ATGCTGAATCTGAGAGAGCACCAGGATTCACAGAGCAGATGAAGCTCCTACAGAAA 764
Qy 601 AGCAGAGCCAGCTGTGCAAGAGAGAGCACCCTGCGGCTGAGCAGAGAGCGCGTCC 660
Db 765 AGCAGAGCCAGCTGTGCAAGAGAGAGCACCCTGCGGCTGAGCAGAGAGCGCGTCC 824
Qy 661 TGGCCCGCAGCAGCTTGAAGAGCCTATGCGGTGAGCTGCAGCGGACCAACCGCTCCCTCA 720
Db 825 TGGCCCGCAGCAGCTTGAAGAGCCTATGCGGTGAGCTGCAGCGGACCAACCGCTCCCTCA 884
Qy 721 AGCAAGAGGTGTGAGCGGCGCCGAGAGAGAGAGAGCGCAAGAGGAGTGAAGTCCG 780
Db 885 AGCAAGAGGTGTGAGCGGCGCCGAGAGAGAGAGAGCGCAAGAGGAGTGAAGTCCG 944
Qy 781 ACTTCCAGGTGACCTGAATGACATTCAGTCCAGATGGAACAGCAGCATGAGCGCACT 840
Db 945 ACTTCCAGGTGACCTGAATGACATTCAGTCCAGATGGAACAGCAGCATGAGCGCACT 1004
Qy 841 CCAAGTGGCCCAAGAGAACATGAGAGTGGCTGAGAGGCTCAAGAGGCTGATTTGAGCAGT 900
Db 1005 CCAAGTGGCCCAAGAGAACATGAGAGTGGCTGAGAGGCTCAAGAGGCTGATTTGAGCAGT 1064
Qy 901 ATGAGTGGCGGAGGAGGATATCGACAAAGTCTTCAACACAGAGGACCTACAGAGAGC 960

QY	481	TCGACGACATTGAATCTCTGAGTACCCACAGAGGAAAGCTGGCTGCTCTGTGTCAAAGAAGT	544
DB	527	TCGACGACATTGAATCTCTGAGTACCCACAGAGGAAAGCTGGCTGCTCTGTGTCAAAGAAGT	586
QY	541	ATGCTGAACTGCTGAGGAGCACCGGAAATTCACAGAAGCAGATGAAGCTCTCTACAGAAA	600
DB	587	ATGCTGAACTGCTGAGGAGCACCGGAAATTCACAGAAGCAGATGAAGCTCTCTACAGAAA	646
QY	601	AGCAGAGCAGCTGCTGCAAGAGAAAGGACCACTTCGCGGTGAGCACAGCAAGGCGCTCC	660
DB	647	AGCAGAGCAGCTGCTGCAAGAGAAAGGACCACTTCGCGGTGAGCACAGCAAGGCGCTCC	706
QY	661	TGGCCCGCAGCAAGCTTCGAGAGCCTATCCGTGAGCTGCAGGGGCAAAACCGCTCCCTCA	720
DB	707	TGGCCCGCAGCAAGCTTCGAGAGCCTATCCGTGAGCTGCAGGGGCAAAACCGCTCCCTCA	766
QY	721	AGGAAGAAAGTGTGCAGCGGGCCCGGAGGAGGAGCAAGCGCAAGAGGTGACCTCGC	780
DB	767	AGGAAGAAAGTGTGCAGCGGGCCCGGAGGAGGAGGAGCAAGCGCAAGAGGTGACCTCGC	826
QY	781	ACTTCCAGGTACACTGAATGACATTCAGCTGCGATGGAACAGCACAATGAGCGCAACT	840
DB	827	ACTTCCAGGTACACTGAATGACATTCAGCTGCGATGGAACAGCACAATGAGCGCAACT	886
QY	841	CAAAGCTCGCAAGAGAAACATGAGCTGGTGTGAGAGGCTCAAGAAGCTGATTGAGCAGT	900
DB	887	CAAAGCTCGCAAGAGAAACATGAGCTGGTGTGAGAGGCTCAAGAAGCTGATTGAGCAGT	946
QY	901	ATGAGTCTCGCGAGGAGCATATCGCAAAAGTCTTCAAAACAAGAGCACTTCAACAGCAGC	960
DB	947	ATGAGTCTCGCGAGGAGCATATCGCAAAAGTCTTCAAAACAAGAGCACTTCAACAGCAGC	1006
QY	961	TGTTGGATGCCAAGCTCCAGCAGGCCACGAGAGATGCTAAGGAGGCGAGAGCGGCAC	1020
DB	1007	TGTTGGATGCCAAGCTCCAGCAGGCCACGAGAGATGCTAAGGAGGCGAGAGCGGCAC	1066
QY	1021	AGCGGAGAAAGGATTTTCTCTGAAAGGCGAGTAGAGTCCACAGAGATGTGTGAGCTGA	1080
DB	1067	AGCGGAGAAAGGATTTTCTCTGAAAGGCGAGTAGAGTCCACAGAGATGTGTGAGCTGA	1126
QY	1081	TGAGCAGCAGAGACACCACTGAGCAACAGCTTGCCCTATACACAGAAAGTTTGAGG	1140
DB	1127	TGAGCAGCAGAGACACCACTGAGCAACAGCTTGCCCTATACACAGAAAGTTTGAGG	1186
QY	1141	AGTTCCAGAACACACTTTCCAAAAGCAGCGAGTATTCCACACATTCACAGCAGAGATGG	1200
DB	1187	AGTTCCAGAACACACTTTCCAAAAGCAGCGAGTATTCCACACATTCACAGCAGAGATGG	1246
QY	1201	AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAAACAACATGTACCGTCCCGGT	1260
DB	1247	AAAAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAAAACAACATGTACCGTCCCGGT	1306
QY	1261	GGGAGAGCAGCAACAGCGCCCTGCTTCAGATGGCTGAGGAGAAACAGTCCCGGATAAG	1320
DB	1307	GGGAGAGCAGCAACAGCGCCCTGCTTCAGATGGCTGAGGAGAAACAGTCCCGGATAAG	1366
QY	1321	AATGAGGGCTGTCAGGTAAAAATCAACCGGTGGAGAAAGCTGTGCGGGCACTGCAGA	1380
DB	1367	AATGAGGGCTGTCAGGTAAAAATCAACCGGTGGAGAAAGCTGTGCGGGCACTGCAGA	1426
QY	1381	CAGAGCGCAATGACCTGAAACAAGAGGCTCAGGACCTGAGTGTGCTGGCCAGGGTCCC	1440
DB	1427	CAGAGCGCAATGACCTGAAACAAGAGGCTCAGGACCTGAGTGTGCTGGCCAGGGTCCC	1486
QY	1441	TCACTGACGTGGCCCTGAGAGGAGCCAGAGGGGCTTGGGGCTCAAGCACCCAGCTCCC	1500
DB	1487	TCACTGACGTGGCCCTGAGAGGAGCCAGAGGGGCTTGGGGCTCAAGCACCCAGCTCCC	1546
QY	1501	CCAGGGTCAAGAGCGCCTTGCTACCCAGGAGCACGAGCAAGAGCATGAGGCCAGA	1560
DB	1547	CCAGGGTCAAGAGCGCCTTGCTACCCAGGAGCACCGAGCAAGAGCATGAGGCCAGA	1606
QY	1561	CTGGGCTCTAAGAGGCCCACTCTCGCCAGGGCC	1592

Db 1607 CTGGGCTCAAGAGCCACCTCCGCCAGGCC 1638

RESULT 8
ABZ34807

ID ABZ34807 standard; cDNA; 1793 BP.

AC ABZ34807;

DT 04-FEB-2003 (first entry)

DE Coding sequence SEQ ID 165, upregulated in osteogenesis.

KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW osteoporosis; bone disease; downregulator; human; ss.

OS Homo sapiens.

PN WO200281745-A2.

PD 17-OCT-2002.

PF 05-APR-2002; 2002WO-IB002211.

PR 05-APR-2001; 2001US-0281400P.

PA (AVET) AVENTIS PHARMA SA.

PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
PI Connolly T, Jackson A, Bushnell SE, Rawadi G;

DR WPI; 2003-058567/05.

PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT useful for bone disease therapy in subject.

PS Claim 26; Page 165-166; 237pp; English.

The present invention relates to novel nucleotide sequences, which are differentially expressed in models of osteogenesis upon being put in contact with a stimulator of osteogenesis. The present sequence is one such sequence. This sequence can be used for diagnosing osteoporosis/bone disease in a patient, promoting osteogenesis and/or preventing osteoporosis/bone disease. The present sequence encodes a secreted protein

Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 27.5%; Score 1294; DB 7; Length 1793;

Best Local Similarity 99.9%; Pred. No. 0;

```
Matches 1414; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

565 GGAATTACAGAAAGCAGATGAAGCTCCTACAGAAAAGCAGAGCCAGCTGGTGCAAGAGA 624

Db
1793 GGAATTACAGAGCAGATGAAGCTCTACAGAAAAGCAGAGCCAGCTGGTGCAAGAGA 1734

625 AGGACCACTGCGCGGTGAGCACAGCAAGGCCGTCTGGCCCGCAGCAAGCTTGAGAGCC 684

Dbb 1733 AGGACCACCTGCGCGGTGAGCACAGCAAGGCCGTCTGGCCCGCAGCAAGCTTGAGAGCC 1674

685 TATGCCGTGAGCTGCAGCGGCAACCGCTCCCTCAAGGAAGGTGTGTGACGCGGCC 744

db
1673 TATGCCGTGAGCTGCAGCGGCACACCGCTCCCTCAAGGAAGGTGTGCAGCGGGCC 1614

745 GGGAGGAGGAGAGCGCAAGGAGGTGACCTCGCACTTCCAGGTGACACTGAAATGACA 804

db 1613 GGGAGGAGGAGAGCGCAAGGAGGTGACCTCGCACTTCCAGGTGACACTGAATGACA 1554

805 TTCAGCTGCAGATGGAACAGCACAAATGAGCGCAACTCCAAGCTGCGCCAAGAGAACATGG 864

1553 TTCAGCTGCAGATGGAAACAGCACAAATGAGCGCAACTCCAAGCTGCGCCAAGAGAACATGG 1494

fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONS reduces side effects. The A-containing ONS break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1880 (AAA32323 to AAA3392) are specifically claimed ONS from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 24.6%; Score 1156; DB 3; Length 1793;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

703 GGCACAAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGCCGCGGAGGAGGAAGC 762
1656 GGCACAAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGCGCCGCGGAGGAGGAAGC 1597
763 GCAGGAGGTGACCTCGCACTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAAC 822
1596 GCAGGAGGTGACCTCGCACTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAAC 1537
823 AGCAATAGCGCAAACTCCAAAGCTGGCGCAAGAGAACTGGAGCTGGCTGAGAGCTCA 882
1536 AGCAATAGCGCAAACTCCAAAGCTGGCGCAAGAGAACTGGAGCTGGCTGAGAGCTCA 1477
883 AGAGCTGATTGACAGTATGAGCTGCGGAGAGGATATCGAAGCTTCGAAACACA 942
1476 AGAGCTGATTGACAGTATGAGCTGCGGAGAGGATATCGAAGCTTCGAAACACA 1417
943 AGGACCTACAAACAGCAGCTGGTGTGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGG 1002
1416 AGGACCTACAAACAGCAGCTGGTGTGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGG 1357
1003 AGGAGAGAGCGGACACAGCGGAGAGAGATTTCTCTGGAAGAGGAGTAGATGCC 1062
1356 AGGAGAGAGCGGACACAGCGGAGAGAGATTTCTCTGGAAGAGGAGTAGATGCC 1297
1063 AGAGATGTGTGAGCTGATGAAGCAGCAGCAGACCCACCTGAGCAACAGCTTGCCCTAT 1122
1296 AGAGATGTGTGAGCTGATGAAGCAGCAGCAGACCCACCTGAGCAACAGCTTGCCCTAT 1237
1123 ACAGAGAGATTTGAGAGTTCCAGACACACTTTCCAAAGCAGGAGGTATTACCA 1182
1236 ACAGAGAGATTTGAGAGTTCCAGACACACTTTCCAAAGCAGGAGGTATTACCA 1177
1183 CATTCAGAGCAGGATGGAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAACCA 1242
1176 CATTCAGAGCAGGATGGAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAACCA 1117
1243 CCATGTACCGTCCCGTGGAGAGCAGCAACAGGCGCTGTGAGATGGCTGAGGAGA 1302
1116 CCATGTACCGTCCCGTGGAGAGCAGCAACAGGCGCTGTGAGATGGCTGAGGAGA 1057
1303 AAACAGTCCCGGATAAAGAACTGGAGGCGCTGAGAGTAAAAATCCAAAGCTGGAGAAC 1362
1056 AAACAGTCCCGGATAAAGAACTGGAGGCGCTGAGAGTAAAAATCCAAAGCTGGAGAAC 997
1363 TGTGCGGCGACTGCAGACAGCGGCAATGACTGACAGAGGGGTACAGGACCTGAGTG 1422
996 TGTGCGGCGCACTGCAGACAGCGGCAATGACTGACAGAGGGGTACAGGACCTGAGTG 937
1423 CTGTGCGGCGGCTCCCTCACTGACAGTGGCGCTGAGAGGAGGCGCAGAGGCGCTGGGG 1482
936 CTGTGCGGCGGCTCCCTCACTGACAGTGGCGCTGAGAGGAGGCGCAGAGGCGCTGGGG 877

QY 1483 CTCAGCACCAGCTCCCGCAGGGTCACAGAGCGCTTGTACCCAGGAGCACCGAGCA 1542
DB 876 CTCAGCACCAGCTCCCGCAGGGTCACAGAGCGCTTGTACCCAGGAGCACCGAGCA 817
QY 1543 CAGAAGCATCAGCCAGACATGGGCGCTCAAGAGCCCACTCCCGCAGGGCTAGAGAGCT 1602
DB 816 CAGAAGCATCAGCCAGACATGGGCGCTCAAGAGCCCACTCCCGCAGGGCTAGAGAGCT 757
QY 1603 GGTCTTGGGTCATGCTGGGAAGGCGCAGCCAGCCAGCGCTGGCCCAATAAAGGCT 1662
DB 756 GGTCTTGGGTCATGCTGGGAAGGCGCAGCCAGCGCGCCAGCCAGCGCTGGCCCAATAAAGGCT 697
QY 1663 CCCATGCTGAGCAGCCCATTTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTGSCA 1722
DB 696 CCCATGCTGAGCAGCCCATTTGCTGAAGCCAGAGATGTTCTTGACCTGGCTGGCATCTGSCA 638
QY 1723 CTTGCAATTTTGGATTTTGTGGGTCAAGTTTACGTACATAGGCGCATTTTGCAGGCGCTTG 1782
DB 637 CTTGCAATTTTGGATTTTGTGGGTCAAGTTTACGTACATAGGCGCATTTTGCAGGCGCTTG 578
QY 1783 CAAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCATTCGGGATGGGGTGTGTACA 1842
DB 577 CAAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCATTCGGGATGGGGTGTGTACA 518
QY 1843 GATGAGTCAAGTGGCTGTGTGAGCTCAAGAGCTTCAGAGGGGCTGTCACTCTGTAGC 1902
DB 517 GATGAGTCAAGTGGCTGTGTGAGCTCAAGAGCTTCAGAGGGGCTGTCACTCTGTAGC 458
QY 1903 TGCCATCAGAGTGGCTGGCAGAGTGTAGCTTGTGATTTCTGTCTGATTTGAGGCTCA 1962
DB 457 TGCCATCAGAGTGGCTGGCAGAGTGTAGCTTGTGATTTCTGTCTGATTTGAGGCTCA 398
QY 1963 GACCCCTCCCTGCCCTT 1979
DB 397 GACCCCTCCCTGCCCTT 381

RESULT 11
AAA35217/c
ID AAA35217 standard; DNA; 1793 BP.
XX
AC AAA35217;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:91.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; anti-inflammatory;
KW antiallergic; antasthmatic; cyostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US017712.
XX
PR 03-AUG-1998; 98US-0095212P.
XX
PA (UPEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX

PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers.

XX
 PS Disclosure; Page 1259; 1343pp; English.

XX The present invention describes a new composition comprising an antisense
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
 CC nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
 CC impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasise to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of the
 CC ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
 CC AAA33992) are specifically claimed ONs from the present invention. N.B.
 CC Sequences given in the disclosure of the present invention do not match
 CC up with their corresponding SEQ ID NO: sequences given in the sequence
 CC listing

XX
 SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 24.6%; Score 1156; DB 3; Length 1793;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	703	GGCACAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGCGCGGAGGAGGAGGAGC	762
DB	1656	GGCACAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGCGCGCGGAGGAGGAGGAGC	1597
QY	763	GCAAGAGGTGACCTCGACATTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAAC	822
DB	1596	GCAAGAGGTGACCTCGACATTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAAC	1537
QY	823	AGCACAATGAGCGCAACTCCAGCTCGCCAGAGAACATGGAGCTGGCTGAGAGGCTCA	882
DB	1536	AGCACAATGAGCGCAACTCCAGCTCGCCAGAGAACATGGAGCTGGCTGAGAGGCTCA	1477
QY	883	AGAACTGATTGAGCAGTATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAAAACA	942
DB	1476	AGAACTGATTGAGCAGTATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAAAACA	1417
QY	943	AGGACCTACACAGCAGCTGGTGGATGCCAGCTCCAGCAGGCCAGGAGATGCTAAGG	1002
DB	1416	AGGACCTACACAGCAGCTGGTGGATGCCAGCTCCAGCAGGCCAGGAGATGCTAAGG	1357
QY	1003	AGGCAGAAGCGGCACCCAGCGGGGAGAAGGATTTCTCTCTGAAAGAGCAGTAGAGTCCC	1062
DB	1356	AGGCAGAAGCGGCACCCAGCGGGGAGAAGGATTTCTCTCTGAAAGAGCAGTAGAGTCCC	1297
QY	1063	AGAGATGTGTGAGCTGATGAGCAGCAGAGAGACCCACCTGAAGCAGCTTGCCTAT	1122
DB	1296	AGAGATGTGTGAGCTGATGAGCAGCAGAGAGACCCACCTGAAGCAGCTTGCCTAT	1237
QY	1123	ACACAGAGAAGTTTTCAGAGTTTCCAGAACACACTTTCAAAAGCAGCGAGTATTACCA	1182
DB	1236	ACACAGAGAAGTTTTCAGAGTTTCCAGAACACACTTTCAAAAGCAGCGAGTATTACCA	1177
QY	1183	CATTCAACAGAGATGAAAAGATGATATGAAGATCAGAAGCTGGAGAAAGAACCA	1242

DB	1176	CATTCAACAGGAGATGGAAGAGTACTAAGAGATCAAGAGCTGGAGAAAGAACCA	1117
QY	1243	CCATGTACCGTCCCGGTGGAGAGCAGCAACAAAGGCCCTGCTGTAGATGGCTGAGAGA	1302
DB	1116	CCATGTACCGTCCCGGTGGAGAGCAGCAACAAAGGCCCTGCTGTAGATGGCTGAGAGA	1057
QY	1303	AAACAGTCCCGGATAAAGAACTGGAGGGCTGCAGAGTAAATCCACGGCTGGAGAAGC	1362
DB	1056	AAACAGTCCCGGATAAAGAACTGGAGGGCTGCAGAGTAAATCCACGGCTGGAGAAGC	997
QY	1363	TGTCGCCGGCAGCTGCACAGAGCGCAATGACCTGAAACAAGAGGGTACAGGACCTGAGT	1422
DB	996	TGTCGCCGGCAGCTGCACAGAGCGCAATGACCTGAAACAAGAGGGTACAGGACCTGAGT	937
QY	1423	CTGTGTCGCGAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCACAGAGGGCTGGG	1482
DB	936	CTGTGTCGCGAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCACAGAGGGCTGGG	877
QY	1483	CTCAAGCACCCAGCTCCCGCCAGGGGTCAACAAGCGCTTGTACCCAGGAGCACCGAGCA	1542
DB	876	CTCAAGCACCCAGCTCCCGCCAGGGGTCAACAAGCGCTTGTACCCAGGAGCACCGAGCA	817
QY	1543	CAGAAGCATCAGGCCAGACTGGGGCTCAAGAGCCCACTCCGCGAGGGCTTAGAGAGCCT	1602
DB	816	CAGAAGCATCAGGCCAGACTGGGGCTCAAGAGCCCACTCCGCGAGGGCTTAGAGAGCCT	757
QY	1603	GGTCTTGGGTGATGCTGGGAAGGAGCGGCGAGCCAGCCAGCGGCTGGCCCATAAAAGCT	1662
DB	756	GGTCTTGGGTGATGCTGGGAAGGAGCGGCGAGCCAGCGGCTGGCCCATAAAAGCT	697
QY	1663	CCCATGCTGAGCAGCCCATTTGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTGCA	1722
DB	696	CCCATGCTGAGCAGCCCATTTGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTGCA	638
QY	1723	CTTGCATTTTGGATTTTGGGTGAGTGTACATAGGGCATTTTCAAGGCTTG	1782
DB	637	CTTGCATTTTGGATTTTGGGTGAGTGTACATAGGGCATTTTCAAGGCTTG	578
QY	1783	CAAAATGATTTTACCTGTAAGTGTACAGTGGGCTTGCATTTGGGGATGGGGGTGTGTACA	1842
DB	577	CAAAATGATTTTACCTGTAAGTGTACAGTGGGCTTGCATTTGGGGATGGGGGTGTGTACA	518
QY	1843	GATGAAGTCACTGGCTTGTCTGTAGCTGAAGTCTTGAAGGGGCTGTCTGTAGC	1902
DB	517	GATGAAGTCACTGGCTTGTCTGTAGCTGAAGTCTTGAAGGGGCTGTCTGTAGC	458
QY	1903	TGCCATCACAGTCAAGTGGCAGAAGTCACTTGAGCATTTCTGTCTGATTTTCAGGCTCA	1962
DB	457	TGCCATCACAGTCAAGTGGCAGAAGTCACTTGAGCATTTCTGTCTGATTTTCAGGCTCA	398
QY	1963	GACCCCTCCCTGCCCTT 1979	
DB	397	GACCCCTCCCTGCCCTT 381	

RESULT 12

AAF21340/c

ID AAF21340 standard; DNA; 1793 BP.

XX

XX AAF21340;

XX

XX 14-MAR-2001 (first entry)

XX

DE Human low adenosine antisense oligonucleotide related sequence #2907.

XX

KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

KW human; airway disorder; bronchoconstriction; lung inflammation;

KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;

KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;

KW respiratory obstruction; pulmonary vasoconstriction; impeded respiration;

KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;

KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;

KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;

chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

WO20062736-A2.

26-OCT-2000.

24-MAR-2000: 2000WO-US008020.

06-APR-1999: 99US-0127958P.

(IVEC-) INTV EAST CAROLINA.

(NYCE/) NYCE J W.

Nyce JW:

WPI: 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions.

Disclosure: Page 1344: 1592pp: English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS) pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 1793 BP: 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 24.6%; Score 1156; DB 3; Length 1793;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1276; Conservative 0; Mismatches 0; Indels 1

703 GGCACAACCGCTCCCTCAAGGAAGGTGTGCAGCGGCCCGGAGGAGGAGGAAGC 762

b
1656 GGCACAACCGCTCCCTCAAGGAAGGTGTGCAGCGGCCGGAGGAGGAGGAGC 1597

763 GCAAGGAGGTGACCTCGCACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAAC 822

1596 GCAAGGAGGTGACCTCGCACCTTCCAGGTGACACTGAATGACATTTCAGCTGCAGATGGAAC 1537

823 AGCACAAATGAGCGCAACTCCAAGCTGCGCCAAAGAGAAACATGGAGCTGGCTGAGAGGCTCA 882

b. 1536 AGCACAAATGAGCGCAACTCCAAGCTGCGCCAAGAGAAATGGAGCTGGCTGAGAGGCTCA 1477

QY	883	AGAAGCTGATTGAGCAGTATGAGCTCCGGAGGAGCATATCGCAAAAGTCTTCAACACA	942
DB	1476	AGAAGCTGATTGAGCAGTATGAGCTCCGGAGGAGCATATCGCAAAAGTCTTCAACACA	1417
QY	943	AGGACCTCAACAACAGCAGCTGGTGGATGCCAAGCTCCAGAGGCCCAAGGAGATGCTAAAGG	1002
DB	1416	AGGACCTCAACAACAGCAGCTGGTGGATGCCAAGCTCCAGAGGCCCAAGGAGATGCTAAAGG	1357
QY	1003	AGGCAGAAAGCGGCGACACAGCGGGGAGAGATTTTCTCTGAAAGAGCGNGTAGAGTCCC	1062
DB	1356	AGGCAGAAAGCGGCGACACAGCGGGGAGAGATTTTCTCTGAAAGAGCGAGTAGAGTCCC	1297
QY	1063	AGAGGATGTGTGAGCTGATGAAGACAGACAAGACCCACCTTGAGCAACAGCTTGCCTCAT	1122
DB	1296	AGAGGATGTGTGAGCTGATGAAGACAGACAAGACCCACCTTGAGCAACAGCTTGCCTCAT	1237
QY	1123	ACACAGAGAAGTTTGAGGAGTTCCAGAAACACACTTCCAAAGCAGCGAGGTATTACCA	1182
DB	1236	ACACAGAGAAGTTTGAGGAGTTCCAGAAACACACTTCCAAAGCAGCGAGGTATTACCA	1177
QY	1183	CATTCAAGCAGGAGATGCGAAAGATCACTAAGAAGATCAAGAACTCGGACAAAGAACCA	1242
DB	1176	CATTCAAGCAGGAGATGCGAAAGATCACTAAGAAGATCAAGAACTCGGACAAAGAACCA	1117
QY	1243	CCATGTACCGGTCCCGTGGGAGAGCAGCAACAAAGGCCCTGCTTGAGATGGCTGAGGAGA	1302
DB	1116	CCATGTACCGGTCCCGTGGGAGAGCAGCAACAAAGGCCCTGCTTGAGATGGCTGAGGAGA	1057
QY	1303	AAACAGTCCGGGATAAAGAACTCGAGGGCTCGAGGTAAAAATCCAAACGGCTCGAGAAAGC	1362
DB	1056	AAACAGTCCGGGATAAAGAACTCGAGGGCTCGAGGTAAAAATCCAAACGGCTCGAGAAAGC	997
QY	1363	TGTGCGGGCACTGACAGACAGAGCGCAATGACCTGAAACAGAGGGTACAGGACCTGAGTG	1422
DB	996	TGTGCGGGCACTGACAGACAGAGCGCAATGACCTGAAACAGAGGGTACAGGACCTGAGTG	937
QY	1423	CTGGTGGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCACAGAGGGCTCGGGG	1482
DB	936	CTGGTGGCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCACAGAGGGCTCGGGG	877
QY	1483	CTCAAGACCCAGCTCCCCAGGGTACAGAGGGCTTGCTTACCCAGGAGCACCAGCA	1542
DB	876	CTCAAGACCCAGCTCCCCAGGGTACAGAGGGCTTGCTTACCCAGGAGCACCAGCA	817
QY	1543	CAGAAGCATCAGGCCAGACTGGCCCTCAAGAGCCCACTCCGCCAGGGCCCTAGAGAGCCT	1602
DB	816	CAGAAGCATCAGGCCAGACTGGCCCTCAAGAGCCCACTCCGCCAGGGCCCTAGAGAGCCT	757
QY	1603	GGTGTGGGTGTCGTGGGAAGGGAGCGGACGCCAGCGCTGGCCCATATAAAGGCT	1662
DB	756	GGTGTGGGTGTCGTGGGAAGGGAGCGGACGCCAGCGCTGGCCCATATAAAGGCT	697
QY	1663	CCCATGCTGAGCAGCCCATTGCTGAGCCAGGATGTC-TGACCTGGCTGGCATCTGGCA	1722
DB	696	CCCATGCTGAGCAGCCCATTGCTGAGCCAGGATGTC-TGACCTGGCTGGCATCTGGCA	638
QY	1723	CTTGCAATTTTGGATTTTGGGTGAGTTTACATAGGGCATTTTGCAGAGGCTTG	1782
DB	637	CTTGCAATTTTGGATTTTGGGTGAGTTTACATAGGGCATTTTGCAGAGGCTTG	578
QY	1783	CAAATGCATTTATACCTGTAAGTGACAGTGGGCTTGCAITGGGGATGGGGGTGTGTACA	1842
DB	577	CAAATGCATTTATACCTGTAAGTGACAGTGGGCTTGCAITGGGGATGGGGGTGTGTACA	518
QY	1843	GATGAAGTCAAGTGGCTTGCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGTAGC	1902
DB	517	GATGAAGTCAAGTGGCTTGCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGTAGC	458
QY	1903	TGCCATCACAGTTCAGTTGGCAGAAAGTACATTTGAGCAITTTCTGTCTGATTTGAGGCTCA	1962
DB	457	TGCCATCACAGTTCAGTTGGCAGAAAGTACATTTGAGCAITTTCTGTCTGATTTGAGGCTCA	398

Db	756	GGTGTGGGTGATCTGTGGAGGAGCGGAGCCAGCGCTGGCCCATAAAGGCT	697
Qy	1663	CCCATGCTGAGCAGCCCATTCGTGAAGCCAGGAGTTCCTTGACCTGGCTGGCATCTGGCA	1722
Db	696	CCCATGCTGAGCAGCCCATTCGTGAAGCCAGGATGTC-TGACCTGGCTGGCATCTGGCA	638
Qy	1723	CTTGCAATTTTGGATTTTGTGGGTCAgTTTTAGTACATAGGCGCATTTGCCAAGGCGCTTG	1782
Db	637	CTTGCAATTTTGGATTTTGTGGGTCAgTTTTAGTACATAGGCGCATTTGCCAAGGCGCTTG	578
Qy	1783	CAAAATGCATTTATACCTGTAAGTGTCAGAGTGGGCTTGCAATTTGGGATGGGGGTGTGTACA	1842
Db	577	CAAAATGCATTTATACCTGTAAGTGTCAGAGTGGGCTTGCAATTTGGGATGGGGGTGTGTACA	518
Qy	1843	GATGAAGTCAGTGGCTTGCTGTGAGCTGAAGAGTCTTTGAGAGGGGCTGTCACTCTGTAGC	1902
Db	517	GATGAAGTCAGTGGCTTGCTGTGAGCTGAAGAGTCTTTGAGAGGGGCTGTCACTCTGTAGC	458
Qy	1903	TGCCATCACAGTCAGTTGGCAGAAGTGACTTGAGCATTTTCTGTCTGTATTTGAGGCTCA	1962
Db	457	TGCCATCACAGTCAGTTGGCAGAAGTGACTTGAGCATTTTCTGTCTGTATTTGAGGCTCA	398
Qy	1963	GACCCCTCCCTGCCCTT	1979
Db	397	GACCCCTCCCTGCCCTT	381

RESULT 14	
ABZ97034/c	
ID ABZ97034	standard; DNA; 1793 BP.
XX	
AC	ABZ97034;
XX	
DT	17-OCT-2003 (first entry)
XX	
DE	Human nucleic acid sequence.
XX	
XX	Human; antitense; lung dysfunction; nasal airway dysfunction;
KW	antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW	antialthematic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW	antitense gene therapy; respiratory; lung; adenosine sensitivity;
KW	adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW	lung inflammation; respiratory disease; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200285308-A2.
XX	
PD	31-OCT-2002.
XX	
PF	23-APR-2002; 2002WO-US013135.
XX	
PR	24-APR-2001; 2001US-0286137P.
XX	
PA	(EPIG-) EPIGENESIS PHARM INC.
XX	
PI	Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI	Miller S, Tang L, Shahabuddin S.
XX	
DR	WPI; 2003-229219/22.
XX	
PT	Pharmaceutical composition for treating ailments associated with impaired
PT	respiration, has oligo(s) antitense to specific gene(s) or its
PT	corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT	ubiquinone.
XX	
XX	Disclosure; SEQ ID NO 12276; 872pp; English.
PS	
XX	
CC	The invention relates to a novel pharmaceutical composition, which has a
CC	first active agent comprising an oligonucleotide antitense to the
CC	initiation codon, coding region, 5' or 3' end genomic flanking regions, of
CC	5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of

	CC	junctions of genes encoding a polypeptide associated with lung and/or
	CC	nasal airway dysfunction and a second active agent comprising an
	CC	antiinflammatory steroid and ubiquinone. A composition of the invention
	CC	has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
	CC	immunosuppressive, and cytostatic activity. The composition may have a
	CC	use in antisense gene therapy. The composition is useful for treating or
	CC	preventing a respiratory, lung or malignant disease or condition, also
	CC	for enhancing the prophylactic or therapeutic respiratory effect of an
	CC	antiinflammatory steroid in a subject, for reducing or depleting levels
	CC	of, or reducing sensitivity to adenosine, reducing levels of adenosine
	CC	receptor, producing bronchodilation, increasing levels of ubiquinone or
	CC	lung surfactant in a subject's tissue, or treating bronchoconstriction,
	CC	lung inflammation, lung allergies, or a respiratory disease or condition.
	CC	Note: The sequence data for this patent is not represented in the printed
	CC	specification, but was obtained in electronic format directly from WIPO
	CC	at ftp.wipo.int/pub/published_pct_sequences
XX		
SQ		Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;
		Query Match 24.6%; Score 1156; DB 7; Length 1793;
		Best Local Similarity 99.9%; Pred. No. 0;
		Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY	703	GGCACAAACCGCTCCCTCAAGGAAGAAGGTGTCACGGGCCCGGAGGAGGAGAAC 762
DB	1656	GGCACAAACCGCTCCCTCAAGGAAGAAGGTGTCACGGGCCCGGAGGAGGAGAAC 1597
QY	763	GMAAGAGGTGACTTCGCACATTCAGGTGACACTGATGACATTTCAGCTGCAGATGGAAC 822
DB	1596	GCAAGAGGTGACTTCGCACATTCAGGTGACACTGATGACATTTCAGCTGCAGATGGAAC 1537
QY	823	AGCACAATGAGCGCAACTCCAAGCTCGCGCAAGACAATGAGGCTGGCTGAGAGCTCA 882
DB	1536	AGCACAATGAGCGCAACTCCAAGCTCGCGCAAGACAATGAGGCTGGCTGAGAGCTCA 1477
QY	883	AGAAGCTGATTGAGCAGTATGAGCTCGCGGAGGAGCATATCCAAAGTTCTTCAAACACA 942
DB	1476	AGAAGCTGATTGAGCAGTATGAGCTCGCGGAGGAGCATATCCAAAGTTCTTCAAACACA 1417
QY	943	AGCACTTCAACAGCAGCTGGTGATGCCAAGCTCCAGCAGGCCACGAGAGATGCTAAAGG 1002
DB	1416	AGCACTTCAACAGCAGCTGGTGATGCCAAGCTCCAGCAGGCCACGAGAGATGCTAAAGG 1357
QY	1003	AGGCAGAAGCGGCACCAGCGGGAAGAGATTTTTCTCTGAAGAGGAGTAGAGTCCC 1062
DB	1356	AGGCAGAAGCGGCACCAGCGGGAAGAGATTTTTCTCTGAAGAGGAGTAGAGTCCC 1297
QY	1063	AGAGGATGTGTGAGCTGATGAGCAGCACAAGAGACCCACCTGAAGCAACAGCTTGCCCTAT 1122
DB	1296	AGAGGATGTGTGAGCTGATGAGCAGCACAAGAGACCCACCTGAAGCAACAGCTTGCCCTAT 1237
QY	1123	ACACAGAGAGTTTGAGGAGTTCAGAAACACATTTCCAAAAGCAGGAGGTATTACCA 1182
DB	1236	ACACAGAGAGTTTGAGGAGTTCAGAAACACATTTCCAAAAGCAGGAGGTATTACCA 1177
QY	1183	CATTCAAGCAGGAGTGGAAAGATGACTATAAGAGATCAAGAGCTGGAGAAAGAACCA 1242
DB	1176	CATTCAAGCAGGAGTGGAAAGATGACTATAAGAGATCAAGAGCTGGAGAAAGAACCA 1117
QY	1243	CCATGTACCCGGTCCCGTGGGAGCAGCAACAAAGCCCTGCTTGAAGTGGCTGAGGAGA 1302
DB	1116	CCATGTACCCGGTCCCGTGGGAGCAGCAACAAAGCCCTGCTTGAAGTGGCTGAGGAGA 1057
QY	1303	AACAGTCCCGGATAAAGAACTGGAGGGCTGCAGGTAAAAATCCAAACGCTGGAGAAGC 1362
DB	1056	AACAGTCCCGGATAAAGAACTGGAGGGCTGCAGGTAAAAATCCAAACGCTGGAGAAGC 997
QY	1363	TGTGCCGGGCACTGCAGACAGACAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTG 1422
DB	996	TGTGCCGGGCACTGCAGACAGACAGCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTG 937
QY	1423	CTGGTGCCCAAGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGCCAGAGGGGCTGGGG 1482

Db 936 CTGGTGGCCAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCGCCAGAGGGGCTGGGG 877
Qy 1483 CTCAGACACCCAGCTCCCGCCAGGGTACAGAGAGCGCTTGTACCCAGGAGCACCGAGCA 1542
Db 876 CTCAGACACCCAGCTCCCGCCAGGGTACAGAGAGCGCTTGTACCCAGGAGCACCGAGCA 817
Qy 1543 CAGAGCATCAGCCAGAGCTGGGCTCAAGAGCCACCTCCGCGAGGCGCTAGAGAGCT 1602
Db 816 CAGAGCATCAGCCAGAGCTGGGCTCAAGAGCCACCTCCGCGAGGCGCTAGAGAGCT 757
Qy 1603 GGTGTGGGTCACTGTGGAGGAGCGGCGAGCCAGAGGCGCTGGGCCATAAAGGCT 1662
Db 756 GGTGTGGGTCACTGTGGAGGAGCGGCGAGCCAGAGGCGCTGGGCCATAAAGGCT 697
Qy 1663 CCCATGCTGAGCAGCCCATTTGTAAGCAGAGATTTTGTACCTGGTGGCATCTGGCA 1722
Db 696 CCCATGCTGAGCAGCCCATTTGTAAGCAGAGATTTTGTACCTGGTGGCATCTGGCA 638
Qy 1723 CTTGCAATTTGGATTTGTGGGTGAGTTTACGTACATAGGCAATTTGCAAGGCTTG 1782
Db 637 CTTGCAATTTGGATTTGTGGGTGAGTTTACGTACATAGGCAATTTGCAAGGCTTG 578
Qy 1783 CAAATGCAATTTATACCTGTAAGTACAGTGGGCTTGCAATGGGCAATGGGGTGTGTACA 1842
Db 577 CAAATGCAATTTATACCTGTAAGTACAGTGGGCTTGCAATGGGCAATGGGGTGTGTACA 518
Qy 1843 GATGAGTCAGTGGCTTGTCTGTAGCTGAGAGTCTTGAGAGGGGCTGTATCTGTAGC 1902
Db 517 GATGAGTCAGTGGCTTGTCTGTAGCTGAGAGTCTTGAGAGGGGCTGTATCTGTAGC 458
Qy 1903 TGCCATCAGTGAAGTGGCAGAGTGAAGTCTGAGCAATTTCTGTCTGATTTGAGGCTCA 1962
Db 457 TGCCATCAGTGAAGTGGCAGAGTGAAGTCTGAGCAATTTCTGTCTGATTTGAGGCTCA 398
Qy 1963 GACCCCTCCCTGCCCTT 1979
Db 397 GACCCCTCCCTGCCCTT 381

RESULT 15

ABZ97033/c
ID ABZ97033 standard; DNA; 1793 BP.
AC ABZ97033;
XX
XX
DT 17-OCT-2003 (first entry)
XX Human-nucleic acid sequence.
DE
XX Homo sapiens.
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW antiallergic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
XX OS
XX Homo sapiens.
XX WO200295308-A2.
XX
XX PN
XX 31-OCT-2002.
XX
XX PD
XX 23-APR-2002; 2002WO-US013135.
XX
XX PF
XX 24-APR-2001; 2001US-0286137P.
XX
XX PR
XX (EPIC-) EPIGENESIS PHARM INC.
XX
XX PA
XX Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
XX PI Miller S, Tang L, Shahabuddin S;
XX PI
XX DR WPI; 2003-229219/22.
XX
XX

PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX

PS Disclosure; SEQ ID NO 12275; 872pp; English.

XX The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiquinone. A composition of the invention
CC has antiinflammatory, antiallergic, antialsthmatic, hypotensive,
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, increasing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Query Match 24.6%; Score 1156; DB 7; Length 1793;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 703 GGCACAAACCCCTCCCTCAAGAGAGAGTGTGAGCGGGCCCGGAGGAGGAGAGC 762
Db 1656 GGCACAAACCCCTCCCTCAAGAGAGAGTGTGAGCGGGCCCGGAGGAGGAGAGC 1597
Qy 763 GCAAGGAGGTGACCTCGACCTCCAGTGACACTGAATGACATTCAGCTCCAGATGGAAC 822
Db 1596 GCAAGGAGGTGACCTCGACCTCCAGTGACACTGAATGACATTCAGCTCCAGATGGAAC 1537
Qy 823 AGCACATGAGCGCAACTCCAAAGCTGGCCAGAGAACATGGAGCTGGCTGAGAGCTCA 882
Db 1536 AGCACATGAGCGCAACTCCAAAGCTGGCCAGAGAACATGGAGCTGGCTGAGAGCTCA 1477
Qy 883 AGAAGCTGATTGAGCAGTATGAGCTGGCGAGGAGCATATCGACAAAGTCTTCARACACA 942
Db 1476 AGAAGCTGATTGAGCAGTATGAGCTGGCGAGGAGCATATCGACAAAGTCTTCARACACA 1417
Qy 943 AGGACCTTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGG 1002
Db 1416 AGGACCTTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGG 1357
Qy 1003 AGGCAGAGAGCGGCGCACCGCGGAGAGGATTTTCTCTGAAAGAGCGAGTAGTCCC 1062
Db 1356 AGGCAGAGAGCGGCGCACCGCGGAGAGGATTTTCTCTGAAAGAGCGAGTAGTCCC 1297
Qy 1063 AGAGGATGCTGAGCTGATGAAGCAGAGACCCACCTTGAAGCAACAGCTTGCCCTAT 1122
Db 1296 AGAGGATGCTGAGCTGATGAAGCAGAGACCCACCTTGAAGCAACAGCTTGCCCTAT 1237
Qy 1123 ACACAGAGAGTGTGAGGAGTTCAGAGACACATTTCCAAAAGACGAGGATTTACCA 1182
Db 1236 ACACAGAGAGTGTGAGGAGTTCAGAGACACATTTCCAAAAGACGAGGATTTACCA 1177
Qy 1183 CATTACAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGATCAAGAGCTGGAGAGAACCA 1242
Db 1176 CATTACAGCAGGAGATGGAAGATGACTAAGAGATCAAGAGCTGGAGAGAACCA 1117
Qy 1243 CCATGTACCCGCTCCCGTGGGAGAGCAGCAACAAAGGCCCTTGCTTGAGATGGCTGAGGAGA 1302
Db 1116 CCATGTACCCGCTCCCGTGGGAGAGCAGCAACAAAGGCCCTTGCTTGAGATGGCTGAGGAGA 1057

Key Location/Qualifiers
 CDS 2100..16932
 FT /*tag= a
 FT /note= "contains introns"
 FT exon 2100..2268
 FT /*tag= b
 FT intron 2269..2996
 FT /*tag= c
 FT exon 2997..3332
 FT /*tag= d
 FT intron 3333..6309
 FT /*tag= e
 FT exon 6310..6401
 FT /*tag= f
 FT intron 6402..9738
 FT /*tag= g
 FT exon 9739..9909
 FT /*tag= h
 FT intron 9910..11839
 FT /*tag= i
 FT exon 11840..12034
 FT /*tag= j
 FT intron 12035..14095
 FT /*tag= k
 FT exon 14096..14215
 FT /*tag= l
 FT intron 14216..14443
 FT /*tag= m
 FT exon 14444..14518
 FT /*tag= n
 FT intron 14519..14944
 FT /*tag= o
 FT exon 14945..15037
 FT /*tag= p
 FT intron 15038..15801
 FT /*tag= q
 FT exon 15802..15897
 FT /*tag= r
 FT intron 15898..16638
 FT /*tag= s
 FT exon 16639..16932
 FT /*tag= t

WO200164874-A2.

07-SEP-2001.

28-FEB-2001; 2001WO-US006356.

02-MAR-2000; 2000US-00517849.

14-JUL-2000; 2000US-00616289.

(BOST-) BOSTON HEART FOUND INC.

Lees AM, Lees RS, Law SW, Arjona AA;

WPI; 2001-565505/63.

P-PSDB; AAB82809.

New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

Example 4; Fig 24; 143pp; English.

The present sequence is that of genomic DNA encoding novel human low density lipoprotein binding protein 3 (LBP-3, see AAB82809). The DNA was isolated from a human genomic library by screening with LBP-3 cDNA (see AAB26501). The open reading frame spans 10 exons. Human LBP-2 nucleic acids are among claimed polynucleotides of the invention that encode novel polypeptides, termed LBPs, capable of binding to native and methylated LDL. Also claimed are isolated LBP polypeptides, and biologically active fragments and analogues of them, as well as expression vectors, cells and methods of producing the LBPs. Methods of

CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 22255 BP; 5195 A; 5302 C; 5924 G; 5834 T; 0 U; 0 Other;

Query Match 18.0%; Score 844; DB 5; Length 22255;
 Best Local Similarity 99.3%; Pred. No. 1.4e-298;
 Matches 2164; Conservative 0; Mismatches 6; Indels 9; Gaps 8;
 YQ 1300 AGAAACAGTCCGGGATAAAGAACTCGAGGGCTCGAGGTAAATAATCCACGGCTGGAGA 1359
 DB 16637 AGAAACAGTCCGGGATAAAGAACTCGAGGGCTCGAGGTAAATAATCCACGGCTGGAGA 16696
 YQ 1360 AGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGA 1419
 DB 16697 AGCTGTCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGGTACAGGACCTGA 16756
 YQ 1420 GTCTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTCAGAGGAGGCCAGAGGGCCCTG 1479
 DB 16757 GTCTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTCAGAGGAGGCCAGAGGGCCCTG 16816
 YQ 1480 GGGCTCAAGCACCCAGCTCCCTCCAGGGTCAAGAGCGCTTGCTACCCAGAGGACCGGA 1539
 DB 16817 GGGCTCAAGCACCCAGCTCCCTCCAGGGTCAAGAGCGCTTGCTACCCAGAGGACCGGA 16876
 YQ 1540 GCACAGAGCATCAGCGCCAGACTGGCCCTCAGAGGCCACCTCCGCCAGGGCCTAGAGAG 1599
 DB 16877 GCACAGAGCATCAGCGCCAGACTGGCCCTCAGAGGCCACCTCCGCCAGGGCCTAGAGAG 16936
 YQ 1600 CCTGTGTGGGTGATGCTGGGAAGGAGCGGCGAGCCAGCCAGGCGCTGGCCATAAAG 1659
 DB 16937 CCTGTGTGGGTGATGCTGGGAAGGAGCGGCGAGCCAGCCAGGCGCTGGCCATAAAG 16996
 YQ 1660 GTCCCATGCTGAGAGCCCATGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTG 1719
 DB 16997 GTCCCATGCTGAGAGCCCATGCTGAAGCCAGGATGTTCTTGACCTGGCTGGCATCTG 17055
 YQ 1720 GCACCTGCAATTTTGGATTGTTGGGTGAGTTTACGTACATAGGGCATTTTGCAGGCG 1779
 DB 17056 GCACCTGCAATTTTGGATTGTTGGGTGAGTTTACGTACATAGGGCATTTTGCAGGCG 17115
 YQ 1780 TTGCAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCATTGGGGATGGGGTGTGT 1839
 DB 17116 TTGCAATGCAATTTATACCTGTAAGTGTACAGTGGGCTTGCATTGGGGATGGGGTGTGT 17175
 YQ 1840 ACAGATCAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGT 1899
 DB 17176 ACAGATCAAGTCAAGTGGCTTGTCTGTGAGCTGAAGAGTCTTGAGAGGGGCTGTCACTGT 17235
 YQ 1900 AGCTGCCATCAGAGTGGTGGCAGAAAGTGAATCTGTCTGTCTGTATTTGAGGC 1959
 DB 17236 AGCTGCCATCAGAGTGGTGGCAGAAAGTGAATCTGTCTGTCTGTATTTGAGGC 17295
 YQ 1960 TCAGACCCCTCCCTGGCCCTTCAGAGCTCAAAACAGTAAATACCAAGTCTTTGACTGC 2019
 DB 17296 TCAGACCCCTCCCTGGCCCTTCAGAGCTCAAAACAGTAAATACCAAGTCTTTGACTGC 17354
 YQ 2020 ATTTGTCTGTGAGCAGGGCTTGTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2078
 DB 17355 ATTTGTCTGTGAGCAGGGCTTGTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 17414
 YQ 2079 TCCTTTGATTTCTAGACCTGGAAAGGTGTCTCTAGGAGAGCCCTGGCAGGGCCCTCA 2138
 DB 17415 TCCTTTGATTTCTAGACCTGGAAAGGTGTCTCTAGGAGAGCCCTGGCAGGGCCCTCA 17474
 YQ 2139 GAGCT-GGGATTTCTCTGGTGGAAACAGGAGCTGGAGATGTTTGTGGTGGGATGATG 2197
 DB 17475 GAGCTGGGATTTGCTGCTTGGAAACAGGAGCTGGAGATGTTTGTGGTGGGATGATG 17534
 YQ 2198 TGCTGTGAGAGCCCTTGGGCATGCTTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2257

18615 CACCAATTTCTCCTTCTCTCTGCCCCAGGCGAGGTACCACTGATCTGATCAAGATTGCGCTG 18678

3333 CTATTCTCTGGTGTATCCTTCCACATCTAGGTGCGCCTCAAGCAGCTGTGTGAGTCTTGAGA 3392

18675 CTATTCTCTGGTGTATCCTTCCACATCTAGGTGCGCCTCAAGCAGCTGTGTGAGTCTTGAGA 18734

3393 TCTCTGCCATCTCTGGCTGAGATACTGCTGTCTCTGTGAGTGTTCCTCATGACCTTTTC 3452

18735 TCTCTGCCATCTCTGGCTGAGATACTGCTGTCTCTGTGAGTGTTCCTCATGACCTTTTC 18794

3453 TTCCCCCTTTGAATCCCTCT 3471

18795 TTCCCCCTTTGAATCCCTCT 18813

RESULT 18
AAQ6987/c
ID AAQ6987 standard; cDNA to mRNA; 1854 BP.

XX AC AAQ6987;
XX AC
XX 25-MAR-2003 (revised)
XX 19-MAR-1995 (first entry)
XX XX
XX Interleukin 14 gene.
XX IL-14; B-cell growth factor; proliferation; differentiation; replication;
XX KW non-Hodgkin's lymphoma; systemic lupus erythematosus;
XX KW systemic necrotising vasculitis; transplant; ss.
XX OS Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX FH 73..1770
XX FT /*tag= a
XX FT
XX FT
XX W09416074-A2.
XX FN
XX 21-JUL-1994.
XX PD
XX 18-JAN-1994; 94WO-US001101.
XX PF
XX 15-JAN-1993; 93US-00005156.
XX PR
(UYWA-) UNIV WASHINGTON ST LOUIS.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (TEXA) UNIV TEXAS SYSTEM.
XX XX
XX PI Ambrus JL, Fauci AS, Ford RJ;
XX XX
XX WPI; 1994-249221/30.
XX DR P-PSDB; AAR55800.
XX XX
XX High molecular weight B cell growth factor - able to stimulate B cell
XX proliferation and inhibit B cell differentiation, useful to treat e.g.
XX systemic lupus erythematosus.
XX Claim 4; Page 75; 95pp; English.

The sequence is that of a gene encoding a high molecular weight B cell growth factor which stimulates B cell proliferation and inhibits B cell differentiation, termed interleukin 14. IL-14 can be used to interfere with the replication of B cells and to treat B cell non-Hodgkin's lymphoma, systemic lupus erythematosus and systemic necrotising vasculitis. (Updated on 25-MAR-2003 to correct PN field.)

XX XX
XX Sequence 1854 BP; 350 A; 559 C; 456 G; 489 T; 0 U; 0 Other;
SQ

Query Match 15.4%; Score 723; DB 2; Length 1854;
Best Local Similarity 99.9%; Pred. No. 3.1e-254;
Matches 773; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

703 GCGCAACCGCTCCCTCAAGGAAGAGTGTGACGCGGCCCGGAGGAGGAGGAGC 762

QY 1141 AGTTCAGACACACTTTCCAAAGCAGCGAGGTATTCACACATTCACGAGGAGATGG 1200
Db 1730 AGTTCAGACACACTTTCCAAAGCAGCGAGGTATTCACACATTCACGAGGAGATGG 1789
QY 1201 AAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGGT 1260
Db 1790 AAAAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGGT 1849
QY 1261 GGGAGAGCAGCAACAGAGCCCTGTTGAGATGGCTGAGAGAGAAACAGTCCGGGATAAG 1320
Db 1850 GGGAGAGCAGCAACAGAGCCCTGTTGAGATGGCTGAGAGAGAAACAGTCCGGGATAAG 1909
QY 1321 AACTGAGGGCTGCGAGGTAAAAATCCAAACGCTGGAGAGCTGTGCGGGCACTGCAGA 1380
Db 1910 AACTGAGGGCTGCGAGGTAAAAATCCAAACGCTGGAGAGCTGTGCGGGCACTGCAGA 1969
QY 1381 CAGAGCGCAATGACCTGAACAGAGG 1406
Db 1970 CAGAGCGCAATGACCTGAACAGAGG 1995
RESULT 20
AAV88391
ID AAV88391 standard; cDNA; 529 BP.
AC AAV88391;
DT 12-FEB-1999 (first entry)
XX EST clone GP232.
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN WC9845437-A2.
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US006956.
PR 10-APR-1997; 97US-00837312.
XX (GEMY) GENETICS INST INC.
PA Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX WPI; 1999-070078/06.
XX New polynucleotides encoding human secreted proteins - derived from e.g.
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.
XX
PS Claim 1; Page 381; 641pp; English.
XX
CC The present 'sequence' represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulatory activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene therapy
XX

SQ Sequence 529 BP; 151 A; 146 C; 168 G; 64 T; 0 U; 0 Other;
Query Match 7.8%; Score 367; DB 2; Length 529;
Best Local Similarity 99.8%; Pred. No. 3.9e-124;
Matches 417; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CAAAAACAGCAGCAGGACCAACCGAGACAGAGCCCGAGGAGCCAGAGCGGCCAGGCC 60
Db 102 CAAAAACAGCAGCAGGACCAACCGAGACAGAGCCCGAGGAGCCAGAGCGGCCAGGCC 161
QY 61 AGCGGGCTCTGCACTAGTAGAGAGAGGTCCCGGACAGCAGCCAGGCTCTCGGAAGCCGG 120
Db 162 AGCGGGCTCTGCACTAGTAGAGAGAGGTCCCGGACAGCAGCCAGGCTCTCGGAAGCCGG 221
QY 121 AGCGGGCTCAAGCCAGAGACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 180
Db 222 AGCGGGCTCAAGCCAGAGACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGA 281
QY 181 GCCGCCCAACTGGAGACATATCTGAGCACAATCTGTGTGGACAATAACAGGGGGGGCCCGC 240
Db 282 GCCGCCCAACTGGAGACATATCTGAGCACAATCTGTGTGGACAATAACAGGGGGGGCCCGC 341
QY 241 GCGAGGATGGGCAACAGGGTGAACCGCTGAACCCGGAAGATGCGAGAAAGTCCCGGACCT 300
Db 342 GCGAGGATGGGCAACAGGGTGAACCGCTGAACCCGGAAGATGCGAGAAAGTCCCGGACCT 401
QY 301 ATGTGCAAGGAATGGGAGCGCTGAACCAACTCCAGTAGTCTATGGAGAGAGAAACCT 360
Db 402 ATGTGCAAGGAATGGGAGCGCTGAACCAACTCCAGTAGTCTATGGAGAGAGAAACCT 461
QY 361 CCAAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGAACGAGGTCCGAGACCGAGA 418
Db 462 CCAAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGAACGAGGTCCGAGACCGAGA 519
RESULT 21
ABK35029
ID ABK35029 standard; cDNA; 1926 BP.
XX
AC ABK35029;
DT 08-MAY-2002 (first entry)
XX
DE Human cDNA encoding secreted protein #167.
KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
KW tissue regeneration; wound healing; burn; haematopoiesis;
KW myeloid cell deficiency; lymphoid cell deficiency.
XX
OS Homo sapiens.
XX
PN WO200177288-A2.
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010224.
XX
PR 06-APR-2000; 2000US-0195582P.
PA (GEMY) GENETICS INST INC.
XX Wong GS, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
XX WPI; 2002-179321/23.
XX
PT Five hundred and ninety two polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for treating

Qy 3695

QY 3695 TGCCTGTGGAGGACTGGAGCTGCTGGATCCACAGTGTAGGAGGCCACAGTGA 375

Db 211 TGCCTGTGGAGGACTGCAGCTGCTGGATCCAGTGTGGTGTAGGAGCCACAGTGA 152
QY 3755 GCAGGTGCCCCCAGCTGGTTCACAGTCAAGGATGTGGCCGCCAGGCAAGGTGCAGCC 3814
Db 151 GCAGGTGCCCCCAGCTGGTTCACAGTCAAGGATGTGGCCGCCAGGCAAGGTGCAGCC 92
QY 3815 TTGTCTCAGCTCCATCCATGTCTAGACCTTC 3847
Db 91 TTGTCTCAGCTCCATCCATGTCTAGACCTTC 59

RESULT 23
ACH39498
ID ACH39498 standard; cDNA; 481 BP.
XX AC ACH39498;
XX XX
DT 13-OCT-2003 (first entry)
XX Human foetal brain cDNA #865.
XX DE
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.
XX XX
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX XX
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 26710; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations.
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 481 BP; 95 A; 133 C; 133 G; 107 T; 0 U; 13 Other;
XX Query Match 5.7%; Score 268; DB 8; Length 481;

Best Local Similarity 100.0%; Pred. No. 5.8e-88;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3575 TTTGCTAGCGAGAGAGGTGCAGGGAAGAGGAGAGGTATGCACCTTCCCTTGAAGAG 3634
Db 214 TTTGCTAGCGAGAGAGGTGCAGGGAAGAGGAGAGGTATGCACCTTCCCTTGAAGAG 273
QY 3635 AGGGGAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACCTCTTAA 3694
Db 274 AGGGGAAGGCTTACAGTGGCCACATAATTGCTGACTCACACTTCAGTACCTCTTAA 333
QY 3695 TGCTGTGAGGAGACTGGAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 3754
Db 334 TGCTGTGAGGAGACTGGAGCTGCTGGATCCAGTGTGGTGTAGGAGGCCACAGTGA 393
QY 3755 GCAGGTGGCCCGAGCTGGTTCACAGTCAAGGATGTGGCCGCCAGGCAAGGTGCAGCC 3814
Db 394 GCAGGTGGCCCGAGCTGGTTCACAGTCAAGGATGTGGCCGCCAGGCAAGGTGCAGCC 453
QY 3815 TTGTCTCAGCTCCATCCATGTCTAGA 3842
Db 454 TTGTCTCAGCTCCATCCATGTCTAGA 481

RESULT 24
ACH24334/c
ID ACH24334 standard; cDNA; 419 BP.
XX AC ACH24334;
XX XX
DT 13-OCT-2003 (first entry)
XX Human adult ovary cDNA #2714.
XX DE
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.
XX XX
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX XX
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 11546; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations.
XX responsible for genetic disorders and other traits. The nucleotide

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (III) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of AAS products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: the sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published](http://wipo.int/pub/published) pct sequences

polypeptide in tissue, as molecular weight markers and as supplement. (II) and its binding partners are useful in method of sites expressing (II). (I) and (II) are useful for treatment involving aberrant protein expression or biological activity polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess and to produce other types of data and products dependent on amino acid sequences. AAS64197-AAS94564 represent novel human coding sequences of the invention. Note: The sequence data patent did not appear in the printed specification, but was electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

of sites expressing (II). (I) and (II) are useful for treating polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess and to produce other types of data and products dependent on amino acid sequences. AAS64197-AAS94564 represent novel human coding sequences of the invention. Note: The sequence data if patent does not appear in the printed specification, but was electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 2626 BP; 717 A; 669 C; 688 G; 551 T; 0 U; 1 Other; Query Match 5.5%; Score 259; DB 5; Length 2626; Best Local Similarity 100.0%; Pred. No. 9.1e-85;

CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: the sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2626 BP; 717 A; 669 C; 688 G; 551 T; 0 U; 1 Other;
Query Match 5.5%; Score 259; DB 5; Length 2626;
Best Local Similarity 100.0%; Pred. No. 9.1e-85;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 3685 TACCTCTTTAAATGCGCTGTGGAGGGAGCTGGAGCTGCTGGATCCCAATGTGGTGGTATAGGAG 374
db 1062 TACCTCTTTAAATGCGCTGTGGAGGGAGCTGGAGCTGCTGGATCCCAATGTGGTGGTATAGGAG 100

CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: the sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 2626 BP; 717 A; 669 C; 688 G; 551 T; 0 U; 1 Other;
	Query Match 5.5%; Score 259; DB 5; Length 2626;
	Best Local Similarity 100.0%; Pred. No. 9.1e-85;
	Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	3695 TACCTCTTAATGCGCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTCTAGGAG 374
DB	1062 TACCTCTTAAATGCGCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGTAGGAG 100
QY	3745 GCCACAGTCAGCAGGTGGCCCGCTGGGTTTCCCAAGTTCAGGAATCTGGCGCCCCAGGCA 380

CC	amino acid sequences.
CC	AA864197-AS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 2626 BP; 717 A; 669 C; 688 G; 551 T; 0 U; 1 Other;
	Query Match 5.5%; Score 259; DB 5; Length 2626;
	Best Local Similarity 100.0%; Pred.No. 9.1e-85;
	Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	3685 TACCTCTTTAATGCCTGTGGAGGACTGGAGTGTGTGATCCCAGTGTTGGTGTAGGAG 374
Db	1062 TACCTCTTTAATGCCTGTGGAGGACTGGAGTGTGTGATCCCAGTGTTGGTGTAGGAG 100
QY	3745 GCCACAGTCAGCAGGTGGCCCCCACCTGGGTTTTCCACGGTCAGAAATCTGGGGCCCCCAGGCA 380
bh	1002 GCCACAGTCAGCAGGTGGCCCCCACCTGGGTTTTCCACGGTCAGAAATCTGGGGCCCCCAGGCA 943

	Sequence	2626 BP;	717 A;	669 C;	688 G;	551 T;	0 U;	1 Other;	
XX	Query Match	5.5%;	Score 259;	DB 5;	Length 2626;				
XX	Best Local Similarity	100.0%;	Pred. No. 9.1e-85;						
XX	Matches 259;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
QY	3685	TACCTCTTTAAATGCCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGGTGTAGGAG	374						374
Db	1062	TACCTCTTTAAATGCCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGGTGTAGGAG	100						100
QY	3745	GCACAGTCAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTCAGAAATGTGGGCCCCCAGGCA	380						380
Db	1002	GCACAGTCAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTCAGAAATGTGGGCCCCCAGGCA	943						943
QY	3805	AGGTGCASCCTTGTCTACAGCTCCATTCATGTCTAGACCTTCAGGGCCAGTCTGTGCAGATG	386						386

	Sequence	2626 BP;	717 A;	669 C;	688 G;	551 T;	0 U;	1 Other;	
XX	Query Match	5.5%;	Score 259;	DB 5;	Length 2626;				
XX	Best Local Similarity	100.0%;	Pred.No. 9.1e-85;						
XX	Matches 259;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
QY	3685	TACCTCTTTAAATGCCCTGTGGAGGGACTGGAGCTGCTGGATCCCATGTCGGTGGTGTAGGAG	374						374
Db	1062	TACCTCTTTAAATGCCCTGTGGAGGGACTGGAGCTGCTGGATCCCATGTCGGTGGTGTAGGAG	100						100
QY	3745	GCACAGTGCAGCGTGGCCCCCAGCTGGGTTTCCACGGTCAGAAATGTGGGCCCCCAGGCA	380						380
Db	1002	GCACAGTGCAGCGTGGCCCCCAGCTGGGTTTCCACGGTCAGAAATGTGGGCCCCCAGGCA	943						943
QY	3805	AGGTGCAGCCCTTGGCTACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCAGATG	386						386
Db	842	AGGTGCAGCCCTTGGCTACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCAGATG	843						843

XX	Sequence	2626 BP	717 A	669 C	688 G	551 T	0 U	1 Other
SQ	Query Match	5.5%	Score 259	DB 5	Length 2626			
	Best Local Similarity	100.0%	Prod. No. 9.1e-85					
	Matches 259	Conservative	0	Mismatches	0	Indels	0	Gaps
QY	3685	TACCTCTTTAAATGCGCTGTGGAGGGACTGGAGTCTGGATCCCACTGTGGTGGTGTAGGAG	374					
Db	1062	TACCTCTTTAAATGCGCTGTGGAGGGACTGGAGTCTGGATCCCACTGTGGTGGTGTAGGAG	100					
QY	3745	GCACAGTGCAGCGTGGCCCCCAGCTGGGTTTCCACGGTCAGAAATGTGGGCCCCCAGGCA	380					
Db	1002	GCACAGTGCAGCGTGGCCCCCAGCTGGGTTTCCACGGTCAGAAATGTGGGCCCCCAGGCA	943					
QY	3805	AGGTGCAGCCCTTGGCTACACGTCCTCCATGCTACAGCTTCAGGCCACGCTCTGCAGATG	386					
Db	942	AGGTGCAGCCCTTGGCTACACGTCCTCCATGCTACAGCTTCAGGCCACGCTCTGCAGATG	893					

	Query Match	5.5%; Score 259; DB 5; Length 2626;
	Best Local Similarity	100.0%; Pred.No. 9.1e-85;
	Matches 259; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	3685	TACCTCTTTAATGCGCTGTGGAGGACTGGAGCTGTGGATCCCACTGTGGTGTAGGAG 374
Db	1062	TACCTCTTTAATGCGCTGTGGAGGACTGGAGCTGTGGATCCCACTGTGGTGTAGGAG 100
QY	3745	GCACAGTCAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTTCAGAAATGTGGGCCCCCAGGCA 380
Db	1002	GCCACAGTCAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTTCAGAAATGTGGGCCCCCAGGCA 943
QY	3805	AGGTGCAGCCTTTGCTCCACAGTCCATTCATGCTACACCTTCAGGCCACTCTGCAGATG 386
Db	942	AGGTGCAGCCTTTGCTCCACAGTCCATTCATGCTACACCTTCAGGCCACTCTGCAGATG 893
QY	3865	AGGTTCCCTACCTTTTCTTCTTCATATGACCAAATTCACCAATGCATCACCTGCTCT 392

	Best Local Match	Local Similarity	Mismatches	Indels	Gaps
QY	3685	TACCTCTTTAATGCCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGTGGAG	0	0	0
Db	1062	TACCTCTTTAATGCCTGTGGAGGACTGGAGCTGCTGGATCCCACTGTGGTGTGGAG	0	0	0
QY	3745	GCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCCAAGTCAGAAATGTGGGCCCCCAGCA	0	0	0
Db	1002	GCCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCCAAGTCAGAAATGTGGGCCCCCAGCA	0	0	0
QY	3805	AGGTGCAGCCTTTCCTCAGCTCCATCTCAGTCTAGACCTTCAGGCCAGCTCTGCAGATG	0	0	0
Db	942	AGGTGCAGCCTTTCCTCAGCTCCATCTCAGTCTAGACCTTCAGGCCAGCTCTGCAGATG	0	0	0
QY	3865	AGGTTCCCTACCTTTTCTTCTTCATTCAGCAAAATCAACCAATCACTACAGCTGCTCT	0	0	0
Db	882	AGGTTCCCTACCTTTTCTTCTTCATTCAGCAAAATCAACCAATCACTACAGCTGCTCT	0	0	0

3685	QY	TACCTCTTTAAATGCTGTGGAGGGA	3744	QY	TCCAGTGTGGTGGTGTAGGAG
		CTGGAGCTGCTGGATCCCACTGTGGTGTAGGAG			
1062	DB	TACCTCTTTAAATGCTGTGGAGGGA	1000	DB	TCCAGTGTGGTGGTGTAGGAG
		CTGGAGCTGCTGGATCCCACTGTGGTGTAGGAG			
3745	QY	GCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCCAAGT	3801	QY	GTGGGCCCCCAAGCA
		CAGAAATGTGGGCCCCCAAGCA			
1002	DB	GCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCCAAGT	943	DB	GTGGGCCCCCAAGCA
		CAGAAATGTGGGCCCCCAAGCA			
3805	QY	AGGTGCAGCCTTTCCTACAGCTCCATCCATGCTACACCTTCAGGCCAGCTCTGCAGATG	3866	QY	CTGCAGATG
		CTGCAGATG			
942	DB	AGGTGCAGCCTTTCCTACAGCTCCATCCATGCTACACCTTCAGGCCAGCTCTGCAGATG	8833	DB	CTGCAGATG
		CTGCAGATG			
3865	QY	AGGTTCCCTACCTTTTCTTCTTCATTCACCAAAATCAACCAATCACTACAGTGTCTCT	3921	QY	CTACAGTGTCTCT
		CTACAGTGTCTCT			
882	DB	AGGTTCCCTACCTTTTCTTCTTCATTCACCAAAATCAACCAATCACTACAGTGTCTCT	8233	DB	CTACAGTGTCTCT
		CTACAGTGTCTCT			
3925	QY	GCCTCTGCTTTTCCAAAGTA	3943	QY	

1062	DB	TACCTCTTTAATGCGCTGTGGAGGACTCGAGCTGCTGGATCCCACTGTGGTGTAGGAG	1000
3745	QY	GCACAGTGCAGCGGTGGCCCCCAGCTGGGTTTCCCAAGTTCAGGAATGTGGSCCCACGGCA	3800
1002	DB	GCACAGTGCAGCGGTGGCCCCCAGCTGGGTTTCCCAAGTTCAGGAATGTGGSCCCACGGCA	943
3805	QY	AGGTGCAGCCTTCTCTACAGCTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATG	3866
942	DB	AGGTGCAGCCTTGTCTACAGCTCCATCCATGCTAGACCTTCAGGCCAGTCTGCAGATG	883
3865	QY	AGGTTCCCTACCTTTTCTTCTTCATATGACCAAAATCAACCAATCACTACAGTGTCTCT	3920
882	DB	AGGTTCCCTACCTTTTCTTCTTCATTCACCAAAATCAACCAATCACTACAGTGTCTCT	8233
3925	QY	GCTTCTGCTTTTCCAAAGTA 3943	
822	DB	GCTTCTGCTTTTCCAAAGTA 804	

3745	QY	GCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCCAGGCA	3800
1002	Db	GCACAGTGAGCAGGTGGCCCCCAGCTGGGTTTCCAGGTGAGGAATGTGGGCCCCCAGGCA	943
3805	QY	AGGTGCAGCCTTTGCTGCACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCGAGATG	3866
942	Db	AGGTGCAGCCTTTGCTGCACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCGAGATG	8833
3865	QY	AGGTTCCCTACCTTTTCTCTTCATATGACAAATCAACCAATCACTACAGCTGCTCT	3922
882	Db	AGGTTCCCTACCTTTTCTCTTCATATGACAAATCAACCAATCACTACAGCTGCTCT	8233
3925	QY	GCTTCTGCTTTCCCAAAGTA 3943	
822	Db	GCTTCTGCTTTCCCAAAGTA 804	

[illegible]

3805	QY	AGGTGCAGCCTTTGTCTCACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCAGATG	386
942	Db	AGGTGCAGCCTTTGTCTCACAGCTCCATCCATGTCTAGACCTTCAGGCCAGTCTGCAGATG	893
3865	QY	AGGTTCCCTACCTTTTCTCTTCATTCAGCCAAATCAACCAATCACTACAGCTGCTCT	392
882	Db	AGGTTCCCTACCTTTTCTCTTCATTCAGCCAAATCAACCAATCACTACAGCTGCTCT	823
3925	QY	GCTTCTGCTTTCCAAAGTA	3943
822	Db	GCTTCTGCTTTCCAAAGTA	804

RESULT 26
AAC05668
ID AAC05668 standard; cDNA; 316 Bp.
XX
AC
AAC05668.

Db	942	AGGTCGAGCCTTTGTCACAGCTCCATGCTAGACCTTCAGGCCAGTCTGCAGATG	893
Qy	3865	AGGTTCCCTACTCTTTTCTCTTTTCATTGACCAAAATCAACCAATCACTACAGCTGCTCT	3922
Db	882	AGGTTCCCTACTCTTTTCTCTCTTTTCATTGACCAAAATCAACCAATCACTACAGCTGCTCT	8233
Qy	3925	GCTTCTGCTTTTCCAAAGTA 3943	
Db	822	GCTTCTGCTTTTCCAAAGTA 804	

RESULT 26
AAC05668
ID AAC05668 standard; cDNA, 316 bp.
XX
AC AAC05668;
XX

3865	AGGTTCCCTACGCTTTTCTTCCTTCATTGACCAAAATCAACCAATCACTACAGCTGCTCT	392
QY		
882	AGGTTCCCTACGCTTTTCTTCCTTCATTGACCAAAATCAACCAATCACTACAGCTGCTCT	823
Db		
3925	GCTTCGTGCTTTCCAAAGTA	3943
QY		
822	GCTTCGTGCTTTCCAAAGTA	804
Db		

RESULT 26
AAC05668
ID AAC05668 standard; cDNA; 316 Bp.
XX
XX AAC05668;
AC
XX
XX
DT 06-OCT-2000 (first entry)
XX
XX

Db	882	AGGTTCCCTACGCTTTTCTTCTTCATTCACCAATCAACCAATCACTACAGCTGCTCT	8233
Qy	3925	GCTTCTGCTTTTCCAAAGTA	3943
Db	822	GCTTCTGCTTTTCCAAAGTA	804

RESULT 26	
AAC05668	
ID ID AAC05668	standard; cDNA; 316 BP.
XX	
AC	AAC05668;
XX	
DT	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein 5' EST, SEQ ID NO: 9743.
XX	
XX	

[illegible]

Db 822 GCTTCGTCTTCCAAAGTA 804

RESULT 26
AAC05668
ID AAC05668 standard; cDNA; 316 BP.
XX
XX AC AAC05668;
XX
XX
DT 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 9743.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens
OS

RESULT 26
AAC05668
ID AAC05668 standard; cDNA; 316 BP.
XX
XX AC
AAC05668;
XX
XX DT 06-OCT-2000 (first entry)
XX
XX DE Human secreted protein 5' EST, SEQ ID NO: 9743.
DE
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX OS Homo sapiens.
XX
XX

RESULT 26	
AAC05668	
ID AAC05668	standard; cDNA, 316 BP.
XX	
XX	
AC	
AC	
XX	AAC05668;
XX	
DT	06-OCT-2000 (first entry)
DT	
XX	
XX	
DE	Human secreted protein 5' EST, SEQ ID NO: 9743.
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
KW	
XX	
XX	Homo sapiens.
OS	
XX	
PN	EP1033401-A2.
XX	
XX	

ID	AAC05668	standard; cDNA; 316 Bp.
XX		
XX	AAC05668;	
AC		
XX		
XX	06-OCT-2000	(first entry)
DT		
XX		
XX	Human secreted protein 5' EST, SEQ ID NO: 9743.	
DE		
XX		
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
KW	gene therapy; chromosome mapping; ss.	
KW		
XX		
XX	Homo sapiens.	
OS		
XX		
XX	EP1033401-A2.	
PN		
XX		
XX	06-SEP-2000.	
PD		
XX		

AC	AAC05668;	
XX		
DT	06-OCT-2000 (first entry)	
XX		
DE	Human secreted protein 5' EST, SEQ ID NO: 9743.	
XX		
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
KW	gene therapy; chromosome mapping; ss.	
XX		
XX	Homo sapiens.	
OS		
XX		
PN	EP1033401-A2.	
XX		
XX		
PD	06-SEP-2000.	
XX		
XX		
PF	21-FEB-2000; 2000EP-00200610.	
XX		
XX		

DT	06-OCT-2000	(first entry)
XX		
XX		
XX	Human secreted protein 5' EST, SEQ ID NO: 9743.	
DE		
DE		
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
KW	gene therapy; chromosome mapping; ss.	
KW		
XX		
XX	Homo sapiens.	
OS		
XX		
XX	EP1033401-A2.	
PN		
XX		
XX		
XX	06-SEP-2000.	
PD		
XX		
XX	21-FEB-2000; 2000EP-00200610.	
PF		
XX		
XX	26-FEB-1999; 99US-0122487P.	
PR		
XX		
XX		

DE	Human secreted protein 5' EST, SEQ ID NO: 9743.
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
XX	
XX	
OS	Homo sapiens.
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-00200610.
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	
XX	

KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX	gene therapy; chromosome mapping; ss.
XX	
XX	Homo sapiens.
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-00200610.
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	

XX OS Homo sapiens.
XX PN EP1033401-A2.
XX XX
XX 06-SEP-2000.
XX PD
XX XX
XX 21-FEB-2000; 2000EP-00200610.
XX PF
XX PR 26-FEB-1999; 99US-0122487P.
XX XX

XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX

XX
XX PD 06-SEP-2000.
XX XX
XX PF 21-FEB-2000; 2000EP-00200610.
XX XX
XX PR 26-FEB-1999; 99US-0132487P.
XX

XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX

XX 26-FEB-1999; 99US-0122487P.
XX
XX

[illegible]

(GEST) GENSET.
Dumas Milne Edwards J, Duclert A, Giordano J;
WPI; 2000-500381/45.
New nucleic acid that is a 5' expressed sequence tag (5' EST) for
obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
diagnostic, forensic, gene therapy and chromosome mapping procedures.
Claim 1; SEQ ID NO 9743; 71pp + Sequence Listing; English.
The present sequence is one of a large number of 5' ESTs derived from
mRNAs encoding secreted proteins. No ORF has yet been conclusively
identified within the present sequence. The 5' ESTs were prepared from
total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
sequences usually correspond mainly to the 3' untranslated region (UTR)
of the mRNA because they are often obtained from oligo-dT primed cDNA
libraries. Such ESTs are not well suited for isolating cDNA sequences
derived from the 5' ends of mRNAs and even in those cases where longer
cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
ESTs are derived from mRNAs with intact 5' ends and can therefore be used
to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
diagnostic, forensic, gene therapy and chromosome mapping procedures.
They are used to obtain upstream regulatory sequences and to design
expression and secretion vectors
SQ Sequence 316 BP; 61 A; 75 C; 83 G; 96 T; 0 U; 1 Other;
Query Match 5.1%; Score 239; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.4e-77; Indels 0; Gaps 0;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2144 GGGATTTCCTGCTGGAAACAAGGACCTGGAGAATGTTTTTCGGTGGATGATGCTGG 2203
DB 78 GGGATTTCCTGCTGGAAACAAGGACCTGGAGAATGTTTTTCGGTGGATGATGCTGG 137
QY 2204 TCAGGAGCCCTTGGGATCGCTTCCCTGCGCTTGGTAGTCCAGGACAGGCCAATG 2263
DB 138 TCAGGAGCCCTTGGGATCGCTTCCCTGCGCTTGGTAGTCCAGGACAGGCCAATG 197
QY 2264 ATGCTTCTCAGTAGCCTTATCATTCACAGGTGCCTCTAGCTGACAAATGATTGACA 2323
DB 198 ATGCTTCTCAGTAGCCTTATCATTCACAGGTGCCTCTAGCTGACAAATGATTGACA 257
QY 2324 AGAGATCAGCAAGGATTTCTCGAAGTGTTTTTCTTATTTCTTTCTTTTCTTTT 2382
DB 258 AGAGATCAGCAAGGATTTCTCGAAGTGTTTTTCTTATTTCTTTCTTTTCTTTT 316
RESULT 27
ABK45487
ID ABK45487 standard; cDNA; 235 BP.
XX AC ABK45487;
XX AC
XX AC
XX DT 05-JUN-2002 (first entry)
XX DE cDNA encoding colon tumour protein, SEQ ID NO 1038.
XX KW Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;
XX KW gene; ss.
XX OS Homo sapiens.
XX PN WO200212328-A2.
XX XX
XX PD 14-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US024218.
XX XX
XX PR 03-AUG-2000; 2000US-0223283P.
XX PR 28-MAR-2001; 2001US-0279763P.
PR 29-JUN-2001; 2001US-0302051P.
XX (CORI-) CORIXA CORP.
XX PI King GE, Meagher MJ, Xu J, Secrist H;
XX WPI; 2002-241739/29.
XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers for
PT the progression of cancer.
XX Claim 1; SEQ ID NO 1038; 147pp; English.
XX The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer. ABK44450-
CC ABK46237 represent coding sequences of human colon tumour proteins of the
CC invention. Note: With the exception of SEQ ID NO 1 and 2, the sequence
CC data for this patent did not form part of the printed specification but
CC was supplied by the European Patent Office
SQ Sequence 235 BP; 43 A; 55 C; 67 G; 70 T; 0 U; 0 Other;
Query Match 5.0%; Score 235; DB 6; Length 235;
Best Local Similarity 100.0%; Pred. No. 7.1e-76; Indels 0; Gaps 0;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 CTGAGAGCTTCATCTCCACCAGATTCTGAGAGTGTCAGCAGCACTTTTTTTTATT 4223
DB 1 CTGAGAGCTTCATCTCCACCAGATTCTGAGAGTGTCAGCAGCACTTTTTTTTATT 60
QY 4224 GTTGTGTTGTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACCTTCGTAGGA 4283
DB 61 GTTGTGTTGTTTCCATGAGTTATCGGACCATGGCTGAGCTCAGGCACCTTCGTAGGA 120
QY 4284 GACTGTGTTATTTCTGTAAGATGGTTATTAACCTCTCCACCCCATCAGGTGGCCCTG 4343
DB 121 GACTGTGTTATTTCTGTAAGATGGTTATTAACCTCTCCACCCCATCAGGTGGCCCTG 180
QY 4344 AGGCTGACCCGAGGCCAGTCGAGCTGCTGTTGTCACGGGGAGGCCAAGG 4398
DB 181 AGGCTGACCCGAGGCCAGTCGAGCTGCTGTTGTCACGGGGAGGCCAAGG 235
RESULT 28
AAS25194/c
ID AAS25194 standard; cDNA; 250 BP.
XX AC AAS25194;
XX AC
XX DT 07-NOV-2001 (first entry)
XX DE Human ovarian PCR-subtracted cDNA library clone #1375.
XX XX
XX KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
XX KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
XX KW primer; probe.
XX OS Homo sapiens.
XX PN WO200157207-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US003733.
XX XX

XX WPI; 1995-206931/27.
 XX Single-stranded DNA for identifying gene signatures - isolated from 3'-
 PT directed human cDNA library that reflects relative abundance of corresp.
 PT mRNA in specific human tissues.
 XX
 XX Claim 1; Page 1478-1479; 2245pp; Japanese.
 XX
 CC A single-stranded DNA (or its complementary strand or the corresp. double
 CC -stranded DNA) which comprises one of the 7837 "GS" sequences given in
 CC AAT19001-T26837 and which is able to hybridize to part of human genomic
 CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were
 CC obtained from 3'-directed cDNA libraries prepared from various human
 CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using
 CC poly(T) as the sole primer. Since the 3'- untranslated sequence is unique
 CC to a particular mRNA species, almost all the 3'-oriented cDNAs hybridize
 CC with specific mRNAs. Each library is constructed so as to reflect
 CC accurately the relative abundance of different mRNAs in the particular
 CC tissue from which it was derived. The appearance frequency of a given GS
 CC in a cDNA library can be determined (esp. using primers and probes
 CC derived from the GS sequences) as a means of diagnosing abnormal cell
 CC function or for recognising different cell types
 XX
 SQ Sequence 268 BP; 51 A; 54 C; 67 G; 82 T; 0 U; 14 Other;
 Query Match 3.5%; Score 166; DB 2; Length 268;
 Best Local Similarity 100.0%; Pred. No. 1.1e-50;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4052 GATCTGTGTGTTTCTTAGATTGCTTCTTCTCCAGGGGACACAGCAGGTGAAGC 4111
 Db 1 GATCTGTGTGTTTCTTAGATTGCTTCTTCTCCAGGGGACACAGCAGGTGAAGC 60
 QY 4112 TCAAGAGCGCATGCTCTGCTTAATAGTAATTTTTCAGGCGCTTGTCCAGCTCAGAGC 4171
 Db 61 TCAAGAGCGCATGCTCTGCTTCTTAATAGTAATTTTTCAGGCGCTTGTCCAGCTCAGAGC 120
 QY 4172 TTCATGTCACAGATTTCTGAGAGGTGTCAGCAGCAGCTTTTTTTT 4217
 Db 121 TTCATGTCACAGATTTCTGAGAGGTGTCAGCAGCAGCTTTTTTTT 166
 RESULT 31
 ID ABL64155
 XX ABL64155 standard; DNA; 441 BP.
 XX
 AC ABL64155;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Stomach cancer related gene sequence SEQ ID NO:2492.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytotoxic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236851P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 XX (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 Soppet DR, Weaver Z;
 WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity; and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 2492; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour

Sequence 441 BP; 85 A; 121 C; 109 G; 124 T; 0 U; 2 Other;

Query Match 3.5%; Score 166; DB 6; Length 441;
 Best Local Similarity 100.0%; Pred. No. 1.1e-50;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3306 ACCACTGATCTGCATGAGTGGCTGCTATTTCTCTGGTGTATCTTCCATCTAGTGC 3365
DB 43 ACCACTGATCTGCATGAGTGGCTGCTATTTCTCTGGTGTATCTTCCATCTAGTGC 102
QY 3366 CCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTCTGCC 3425
DB 103 CCTCAAGCAGCTGTGTGAGTGTGAGATCTCTGCCATCTCTGGCTGAGATCTGTCTGCC 162
QY 3426 TGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTGAATCCCTCT 3471
DB 163 TGTGAAGTGTTCCTCCATGACCTTTTCTTCCCTTGAATCCCTCT 208

RESULT 32

ADD49735/c

ID ADD49735 standard; cDNA; 236 BP.

XX AC ADD49735;

XX DT 15-JAN-2004 (first entry)

XX DE Human lung cancer associated cDNA 61594712.

XX KW Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;
vaccine; T-cell; tumour.

XX OS Homo sapiens.

XX PN US2003194764-A1.

XX PD 16-OCT-2003.

XX PF 04-APR-2002; 2002US-00116712.

XX PR 05-APR-2001; 2001US-0282289P.

XX PS 05-OCT-2001; 2001US-0327511P.

XX PA (CORI-) CORIXA CORP.

XX PI Bangor CS, Switzer A;

XX PR WPI; 2003-844452/78.

XX PT New isolated polypeptides and polynucleotides useful for diagnosing,
preventing and treating cancer, particularly lung cancer.

XX PS Claim 1; SEQ ID NO 467; 250pp; English.

XX CC The invention relates to an isolated polynucleotide (a) comprising any of
the 666 fully defined nucleotide sequences appearing as ADD49269 -
ADD49335, ADD49336 and ADD49338, complements of (a); sequences of at
least 20 contiguous residues of (a); sequences that hybridise to (a)
under highly stringent conditions; sequences having at least 75 or 90%
identity to (a); or degenerate variants of (a). Also included are an
isolated polypeptide (b) (comprising: sequences encoded by the new
polynucleotide; any of the 4 amino acid sequences fully defined in the
specification; or sequences having at least 70 or 90% identity to the
sequence in (a) or (b)), an expression vector comprising the above
polynucleotide operably linked to an expression control sequence, a host
cell transformed or transfected with the above expression vector, an
isolated antibody, or its antigen-binding fragment, that specifically
binds to the above polypeptide, an oligonucleotide that hybridises to the
above-mentioned nucleotide sequences under highly stringent conditions, a
fusion protein comprising at least one polypeptide cited above, detecting
the presence of a cancer in a patient (comprising: obtaining a biological
sample from the patient; contacting the biological sample with a binding
agent that binds to the polypeptide, or with the oligonucleotide cited
above; detecting in the sample an amount of the polypeptide that binds to
the binding agent, or an amount of a polynucleotide that hybridises to
the oligonucleotide; and comparing the amount of polypeptide, or
polynucleotide that hybridises to the oligonucleotide, to a predetermined
cut-off value and then determining the presence of a cancer in the
patient), a method for stimulating and/or expanding T-cells specific for

CC a tumour protein (comprising contacting T-cells with the above
polypeptide, polynucleotide or antigen-presenting cells that express the
polynucleotide, under conditions and for a time sufficient to permit the
stimulation and/or expansion of T-cells), an isolated T-cell population
comprising T-cells prepared by the method, a composition comprising a
first component selected from physiological carriers and
immunostimulants, and a second component selected from the above
polypeptide, polynucleotide, antibody, fusion protein, T-cell population
and antigen-presenting cells that express the above polypeptide,
stimulating an immune response in a patient (comprising administering to
the patient the above composition) treating lung cancer in a patient
(comprising administering to the patient the above composition and a
diagnostic kit (comprising: at least one oligonucleotide cited above; or
at least one antibody cited above and a detection reagent, where the
detection reagent comprises a reporter group). The composition and
methods are useful in diagnosing, preventing and treating cancer,
particularly lung cancer. The present sequence is a lung cancer-
associated antigen cDNA of the invention.

XX SQ Sequence 236 BP; 69 A; 67 C; 55 G; 45 T; 0 U; 0 Other;

Query Match 3.3%; Score 153; DB 9; Length 236;

Best Local Similarity 100.0%; Pred. NO. 6.4e-46;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4210 TTTTCTTTTATTTGTTGTTGTTTCCATGAGGTTATCGACCATGGGCTGAGCTCAGG 4269

DB 191 TTTTCTTTTATTTGTTGTTGTTTCCATGAGGTTATCGACCATGGGCTGAGCTCAGG 132

QY 4270 CACTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTTAAACCTCTCCACCCCA 4329

DB 131 CACTTCTGTAGGAGACTGTTATTTCTGTAAGATGTTTAAACCTCTCCACCCCA 72

QY 4330 TCACGGTGGCCCTGAGGCTGACCCGGAGGCCA 4362

DB 71 TCACGGTGGCCCTGAGGCTGACCCGGAGGCCA 39

RESULT 33

ACH49990

ID ACH49990 standard; cDNA; 457 BP.

XX AC ACH49990;

XX DT 13-OCT-2003 (first entry)

XX DE Human leukocyte cDNA #1584.

XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

XX PA (DRMA/) DRMANAC R T.

XX PA (LABA/) LABAT I.

XX PA (STAC/) STACHE-CRAIN B.

XX PA (DICK/) DICKSON M C.

XX PA (JONE/) JONES L W.

XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX DR WPI; 2003-615964/58.

XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
as hybridization probes, as oligomers for PCR, for chromosome and gene

PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX
PS Claim 1; SEQ ID NO 37202; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH2789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 457 BP; 106 A; 109 C; 94 G; 148 T; 0 U; 0 Other;
Query Match 3.2%; Score 152; DB 8; Length 457;
Best Local Similarity 99.5%; Pred. No. 1.4e-45;
Matches 202; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2834 GCCATAGCTGTTGTTCTAACAGTAAGAATGAATTAAGTGAATGACCTGGCCCA 2893
Db 61 GCCATAGCTGTTGTTCTAACAGTAAGAATGAATTAAGTGAATGACCTGGCCCA 120
QY 2894 ATCCATAAATGGCTGCACACAGTGTCTTCTGTCCTCTTCTACCCCACTCCAGTACA 2953
Db 121 ATCCATAAATGGCTGCACACAGTGTCTTCTGTCCTCTTCTACCCCACTCCAGTACA 180
QY 2954 TAACTACTATGACTGTGTAGAGCCATTCTATATGCTGAATGTTCTGCTGTTGCAACTT 3013
Db 181 TAACTACTATGACTGTGTAGAGCCATTCTATATGCTGAATGTTCTGCTGTTGCAACTT 240
QY 3014 GCCAGGTTATAGCCAGTGTGTTG 3036
Db 241 GCCAGGTTATAGCCAGTGTGTTG 263
RESULT 34
ID AAS65097/c
AC AAS65097 standard; cDNA; 473 BP.
XX
XX AAS65097;
XX
XX 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #901.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;

XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG00910.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 901; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 473 BP; 135 A; 108 C; 95 G; 135 T; 0 U; 0 Other;
Query Match 3.2%; Score 150; DB 5; Length 473;
Best Local Similarity 100.0%; Pred. No. 7.3e-45;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2834 GCCATAGCTGTTGTTCTAACAGTAAGAATGAATTAAGTGAATGACCTGGCCCA 2893
Db 437 GCCATAGCTGTTGTTCTAACAGTAAGAATGAATTAAGTGAATGACCTGGCCCA 378
QY 2894 ATCCATAAATGGCTGCACAGTGTGTTCTGTCCTCTTCTACCCCACTCCAGTACA 2953
Db 377 ATCCATAAATGGCTGCACAGTGTGTTCTGTCCTCTTCTACCCCACTCCAGTACA 318
QY 2954 TAACTACTATGACTGTGTAGAGCCATTCT 2983
Db 317 TAACTACTATGACTGTGTAGAGCCATTCT 288
RESULT 35
ID ACH23051 standard; cDNA; 439 BP.
XX
XX ACH23051;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human adult ovary cDNA #1431.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX


```
21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246509P.
PR 08-NOV-2000; 2000US-0246510P.
PR 08-NOV-2000; 2000US-0246511P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0248527P.
PR 08-NOV-2000; 2000US-0248528P.
PR 08-NOV-2000; 2000US-0248532P.
PR 08-NOV-2000; 2000US-0248609P.
PR 08-NOV-2000; 2000US-0248610P.
PR 08-NOV-2000; 2000US-0248611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451930/48.
DR P-PSDB; AAU22308.
XX
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system.
XX
PS: Claim 1; SEQ ID NO 467; 674pp; English.
XX
CC Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the
CC cardiovascular system antigen polypeptides of the invention.
CC Cardiovascular system antigens and their associated polynucleotides are
CC useful in the diagnosis, treatment and prevention of various types of
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. A pathological condition can be determined by
CC detecting the presence or absence of a mutation in a cardiovascular
CC system antigen polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma and
CC pleurisy. The polypeptides can also be used to aid wound healing, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues and in chemotaxis. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences

Query Match 3.0%; Score 140; DB 4; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.4e-41;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 184 GCCAACTGGGAAGACATACCTAGGACATACCTGTGGACAAATACCGAGGGGGCCCCGGCG 243
DB 88 GCCAACTGGGAAGACATACCTAGGACATACCTGTGGACAAATACCGAGGGGGCCCCGGCG 147
OY 244 AGGATGGGGGCACAGGGTGAGCCGGCTGAACCCGGAAGATGCAGAGAACTCCCGACCTATG 303
DB 148 AGGATGGGGGCACAGGGTGAGCCGGCTGAACCCGGAAGATGCAGAGAACTCCCGACCTATG 207
OY 304 TGGCAAGGAATGGGAGCCT 323
DB 208 TGGCAAGGAATGGGAGCCT 227

RESULT 37
ADE45661
ID ADE45661 standard; cDNA; 405 BP.
XX
AC ADE45661;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human cardiovascular system related polynucleotide #457.
XX
KW Human; cardiovascular system related polypeptide; cancer;
KW proliferative disorder; foetal abnormality; developmental abnormality;
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
```


PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764869.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Barash SC;
 XX PI WPI; 2003-743766/70.
 XX DR P-PSDB; ADE46276.
 XX New cardiovascular system related polynucleotides and polypeptides,
 PT useful for preventing, treating, or ameliorating a medical condition,
 PT such as cancer of cardiovascular tissues and cancer metastases.
 XX Claim 1; SEQ ID NO 467; 262pp; English.
 XX The invention relates to human cardiovascular system related polypeptides
 CC and the polynucleotides encoding them. The polypeptides, polynucleotides
 CC and antibodies to the polypeptides are useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition,
 CC for preventing, treating, or ameliorating a medical condition, such as
 CC cancer of cardiovascular system tissues, proliferative disorders, foetal
 CC and developmental abnormalities, haematopoietic disorders, diseases of
 CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
 CC arthritis), inflammation, allergies, neurological disorders (e.g.,
 CC Alzheimer's disease, Parkinson's disease), cognitive disorders,
 CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
 CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic
 CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
 CC related disorders, endocrine disorders and infections. The nucleic acids
 CC are also useful for chromosome identification, radiation hybrid mapping
 CC or long-range restriction mapping. The polypeptides and polynucleotides
 CC may also be used as food additives or preservatives to increase or
 CC decrease storage capabilities, fat content or other nutritional
 CC components. This sequence represents a human cardiovascular system
 CC related polynucleotide of the invention.
 XX SQ Sequence 405 BP; 123 A; 90 C; 126 G; 63 T; 0 U; 3 Other;
 Query Match 3.0%; Score 140; DB 9; Length 405;
 Best Local Similarity 100.0%; Pred. No. 3.4e-41;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1
 QY 184 GCAACTGGAAGACATCTAGACACATCTGTGTGGACAAATACCGAGGGGCCCCGGCG 243
 Db 88 GCAACTGGAAGACATCTAGACACATCTGTGTGGACAAATACCGAGGGGCCCCGGCG 147
 QY 244 AGGATGGGCGACAGGGTGGAGCGGCTGACCCGAGATCGACAGAGTCCCGACCTATG 303
 Db 148 AGATGGGCGACAGGGTGGAGCGGCTGACCCGAGATCGACAGAGTCCCGACCTATG 207
 QY 304 TGGCAAGGAATGGGAGCCT 323
 Db 208 TGGCAAGGAATGGGAGCCT 227
 RESULT 38
 ACH27023
 ID ACH27023 standard; cDNA; 416 BP.
 XX ACH27023;
 XX ACH27023;
 XX 13-OCT-2003 (first entry)
 XX Human adult ovary cDNA #5403.
 XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 XX Homo sapiens.
 OS US2003073623-A1.
 PN

XX 17-APR-2003.
 XX 30-JUL-2001; 2001US-00918995.
 XX 30-JUL-2001; 2001US-00918995.
 XX (DRNA/) DRMANAC R T.
 XX (LABA/) LABAT I.
 XX (STAC/) STACHE-CRAIN B.
 XX (DICK/) DICKSON M C.
 XX (JONE/) JONES L W.
 XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX WPI; 2003-615964/58.
 XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX Claim 1; SEQ ID NO 14235; 44pp; English.
 XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623
 XX SQ Sequence 416 BP; 141 A; 64 C; 54 G; 67 T; 0 U; 90 Other;
 Query Match 2.7%; Score 129; DB 8; Length 416;
 Best Local Similarity 99.1%; Pred. No. 3.5e-37;
 Matches 229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4450 CACGAGGTGGTCTACCCCGAGGACGACGACGACCTGCTCTTGGCCCTTCTGGGGGA 4509
 Db 1 CACGAGGTGGTCTACCCCGAGGACGACGACGACCTGCTCTTGGCCCTTCTGGGGGA 60
 QY 4510 AAGGAGCTGCCCTTCTGTCCTGTAACTGCTTTTCCCTATGCCCCAACCCGCCACTCAGAC 4569
 Db 61 AGGAGCTGCCCTTCTGTCCTGTAACTGCTTTTCCCTATGCCCCAACCCGCCACTCAGAC 120
 QY 4570 TTGTTGAAGCTGCACCTGGCAGCTTTTGTCTCTTTGGGTATTCAACAGCAGGGA 4629
 Db 121 TTGTTGAAGCTGCACCTGGCAGCTTTTGTCTCTTTGGGTATTCAACAGCAGGGA 180
 QY 4630 CTTCATTTTGATGTATTTTAAACACATTAATAAGAGTCTGTGCTTAA 4680
 Db 181 CTTCATTTTGATGTATTTTAAACACATTAATAAGAGTCTGTGCTTAA 231
 RESULT 39
 ACH23800
 ID ACH23800 standard; cDNA; 547 BP.
 XX ACH23800;
 XX ACH23800;
 XX 13-OCT-2003 (first entry)
 XX

	RESULT 40	
XX	ABV97036	ID ID ABV97036 standard; cDNA; 264 BP.
XX	AC AC	ABV97036;
XX	DD DT	14-JAN-2003 (first entry)
XX	EE EE	Human pancreatic cancer expressed cDNA SEQ ID NO 2444.
XX	KW KW	Human; pancreas; cancer; gene therapy; vaccine; immunostimulant; cytotoxic; tumour; gene; ss.
XX	DOS DOS	Homo sapiens.
XX	NN NN	WO200260317-A2.
XX	PED PED	08-AUG-2002.
XX	FF FF	30-JAN-2002; 2002MO-USO02781.
XX	PP PP	30-JAN-2001; 2001US-026530SP.
XX	RPR RPR	31-JAN-2001; 2001US-026568BP.
XX	FEB FEB	09-FEB-2001; 2001US-026756BP.
XX	MAR MAR	21-MAR-2001; 2001US-0278651P.
XX	APR APR	28-APR-2001; 2001US-0287112P.
XX	MAY MAY	16-MAY-2001; 2001US-0291631P.
XX	JUL JUL	12-JUL-2001; 2001US-0305484P.
XX	AUG AUG	20-AUG-2001; 2001US-0313999P.
XX	NOV NOV	27-NOV-2001; 2001US-0333626P.
XX	COR COR	(CORI-) CORIXA CORP.
XX	BEN BEN	Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
PI PI	DR DR	WPI; 2002-627435/67.
XX	PT PT	New isolated polynucleotide and pancreatic tumor polypeptides, useful for diagnosing, preventing and/or treating cancer, particularly pancreatic cancer.
XX	PS PS	Claim 1; SEQ ID NO 2444; 300pp + Sequence Listing; English.
CC CC	CCC CCC	The invention relates to an isolated polynucleotide (I) comprising: (a) any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b) complements of (a); (c) sequences consisting of at least 20 contiguous residues of (a); (d) sequences that hybridize to (a), under moderately stringent conditions; (e) sequences having at least 75% or 90% identity to (a); or (f) degenerate variants of (a). Polypeptides (ASP68596-
CC CC	CCC CCC	ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer in a patient and compositions comprising polypeptides, polynucleotides, antibodies, fusion proteins. T cell populations and antigen presenting cells expressing the polypeptide are useful in treating pancreatic cancer. probes or primers for immune response. The polynucleotides can be used as and stimulating an immune response. In the design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in the tumour cells, in vaccines and for gene therapy. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC CC	XXX XXX	Sequence 264 BP; 61 A; 56 C; 56 G; 91 T; 0 U; 0 Other;
XQ XQ	Match Match	Query Match 2.3%; Score 108; DB 6; Length 264;
	Best Local Similarity	Best Local similarity 99.4%; Pred. No. 1.7e-29;
	Matches 158; Conservative	Matches 158; Mismatches 1; Indels 0; Gaps 0;
OY OY	2966 ACTGTGAGACCATTCTATGCCTGAATGGTTCTCCTTTCACAACITGCCAGCGTATTA	3025
DB DB		ACTGTGAGACCATTCTATGCCTGAATGGTTCTCCTTTCACAACITGCCAGCGTATTA 60

[illegible]

RESULT 41
AAA45144
ID AAA45144 standard; cDNA; 245 bp.

AAA
AC
XX
DT

AAA45144;
21-AUG-2000 (first entry)

Human secreted expressed sequence tag SEQ ID NO:1719.

Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
expressed sequence tag; EST; probe; chemotactic; proliferative;
immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;
antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;
cerebroprotective; anticoagulant; antidepressant; gene therapy; vaccine;
autoimmune disorder; multiple sclerosis; allergic condition;
insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
central nervous system disorder; Alzheimer's disease; stroke;
Parkinson's disease; Huntington's disease; coagulation disorder;
haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
infection; depression; psoriasis; ss.

PN WO200021991-A1.

PD 20-APR-2000.

15-OCT-1999; 99WO-US024206.

PR 15-OCT-1998; 98US-0104436P.

PA (GEMY) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
Merberg D; Treacy M, Bowman MR;

AA
DR
WPI; 2000-317938/27.

Isolated polynucleotides, and encoded proteins, comprising secreted PT expressed sequence tags (ESTs), useful for treating various disorders such as autoimmune, infectious, and central nervous system disorders.

PS Claim 1; Page 645; 803pp; English.

AAA43426 to AAA45925 represent specifically claimed secreted expressed sequence tags (SESTs), isolated from human, mouse, chicken and rat tissue sources. The SESTs can have a range of activities depending on the tissues they were isolated from. The activities include: chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal; antiviral; antididiabetic; antiaesthatic; vulnerrary; antiulcer; osteoprotective; neuroprotective; nootropic; antiparkinsonian; antispasmodic; cerebroprotective; anticonvulsant; and antidepressant. The SESTs can be used for gene therapy and in vaccines. The SESTs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the SESTs. Proteins encoded by the SESTs are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, osteoporosis, osteoarthritis, central nervous system disorders (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation

CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention
xx
SQ Sequence 245 BP; 43 A; 56 C; 53 G; 93 T; 0 U; 0 Other;

```
Query Match      1.5%; Score 71; DB 3; Length 245;
...Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 71; Conservative 0; Mismatches 0; Indels
```

Qy	2292	GGTGCCTCTTAGCCTGCACAAATGATTGACAAGAGATCACCCAAAGGATTATTCCTGAA	2351
Db	21	GSTGCCTCTTAGCCTGCACAAATGATTGACAAGAGATCACCCAAAGGATTATTCCTGAA	80

Qy 2352 GGTGTTTTTT 2362
pb 81 GGTGTTTTTT 91

RESULT 42
AAV32836
ID AAV32836 standard: cdNA: 4722 BP.

AAV32836;

DT 09-NOV-1998 (first entry)

DE Rabbit low density lipoprotein binding protein LBP-3 cDNA.

Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW
KW
KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine; ss.

OS Oryctolagus cuniculus.

id	Key	Location/Qualifiers
PH	CDS	61..1734
FT		/*tag= a
FT		348..390
FT	misc_feature	/*tag= b
FT		/*note= "Claim 22"

PN WO9823282-A1.

04-JUN-1998.

PF 26-NOV-1997; 97WO-US021857.

PR 27-NOV-1996; 96US-0031930P.

XX
YF 09-NOV-68 17:03-004024/E

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

XX
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840

DR P-PSDB; AAW49039.

PT Nucleic acid encoding low density lipoprotein binding proteins and
PT related vectors - transformed cells, proteins, and modulators of binding,
PT useful for treatment and diagnosis of atherosclerosis and for identifying
PT subjects at risk.

PS Claim 7: Fig 14: 47pp: English.

This cDNA clone codes for novel rabbit low density lipoprotein (LDL) binding protein LBP-3 (see AAV49039). It was isolated by functional screening of a cDNA library, produced from balloon-deendothelialised healing rabbit abdominal aorta, mRNA, for clones encoding LBPs able to bind both native and methyl LDL. cDNA clones (see AAV32834-39), and encoded rabbit and human LBPs (see AAV49037-42) are claimed. An abnormality in an aspect of LBP metabolism or structure is diagnostic for atherosclerosis. The invention provides methods for a risk for atherosclerosis.

CC determining if an animal is at risk for atherosclerosis (e.g. for
 CC prenatal screening); methods for treating atherosclerosis (including gene
 CC therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent
 CC formation of atherosclerotic plaque; and methods for treating a cell
 CC having an abnormality in LBP structure or metabolism. Pharmaceutical and
 CC vaccine compositions are also provided, as well as recombinant vectors
 CC and host cells used to produce recombinant LBP

XX Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;

Query Match 1.3%; Score 59; DB 2; Length 4722;
 Best Local Similarity 100.0%; Pred. No. 9.6e-12; Indels 0; Gaps 0;
 Matches 59; Conservative 0; Mismatches 0;

QY 1350 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGTT 1408
 DB 1456 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGTT 1514

RESULT 43

AAH26491
 ID AAH26491 standard; cDNA; 4722 BP.

AC AAH26491;

DT 12-NOV-2001 (first entry)

XX Rabbit low density lipoprotein binding protein 3 (LBP-3) cDNA.

XX Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
 KW ss.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers
 FH 25..1734
 FT CDS /*tag= a

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX P-PSDB; AAB82801.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 5; Fig 14; 143pp; English.

XX The present sequence is that of cDNA encoding novel rabbit low density
 CC lipoprotein binding protein 3 (LBP-3, see AAB82801). The cDNA was
 CC isolated following screening of a rabbit cDNA library for clones encoding
 CC LBP3 that bound to both native low density lipoprotein (LDL) and methyl
 CC LDL. The invention provides claimed polynucleotides encoding novel
 CC polypeptides which are capable of binding to native and methylated LDL,
 CC and analogues of them, as well as expression vectors, cells and methods
 CC of producing the LBP3. Also claimed are methods of determining if an
 CC animal is at risk for atherosclerosis, methods for evaluating an agent
 CC for use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP. Pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions, are also claimed

SQ Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;

Query Match 1.3%; Score 59; DB 5; Length 4722;
 Best Local Similarity 100.0%; Pred. No. 9.6e-12;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1350 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGTT 1408
 DB 1456 CGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGGTT 1514

RESULT 44

ABN74519

ID ABN74519 standard; cDNA; 937 BP.

XX AC ABN74519;

XX 03-JUL-2002 (first entry)

XX Bovine embryonic germ (EG) cell cDNA EST #570.

XX Bovine; Bos taurus; EST; expressed sequence tag; totipotence;
 KW development; gene; ss.

XX Bos taurus.

XX WO200194550-A2.

XX 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US018576.

XX 07-JUN-2000; 2000US-0209874P.

XX 06-JUN-2001; 2001US-00876143.

XX (INFI-) INFIGEN INC.

XX Eilertsen KJ, Pfister-Genskow M, Childs L;

XX WPI; 2002-351289/38.

XX An expressed sequence tag (EST), the expression of which, or its
 PT complementary sequence, in a cell identifies the cell as a
 PT developmentally competent or incompetent cell.

XX Example 16; Page 461-462; 584pp; English.

XX The present invention describes an expressed sequence tag (EST), where
 CC the EST is an isolated, enriched, or purified nucleic acid sequence
 CC representing all or part of a gene, the expression of which, or its
 CC complementary sequence, in a cell identifies the cell as a
 CC developmentally competent or incompetent cell. Molecules which induce
 CC developmental competence in a cell line are useful for inducing
 CC totipotence in one or more cells. Molecules which induce developmental
 CC totipotence in a cell line are useful for preventing a full term
 CC pregnancy in an animal and inhibiting totipotence. The molecules are also
 CC useful for treating a disease in an animal by inducing development of one
 CC or more cells of the animal into a specific cell type. The present
 CC sequence represents a bovine EST which is given in the exemplification of
 CC the present invention

SQ Sequence 937 BP; 222 A; 211 C; 265 G; 203 T; 0 U; 36 Other;

Query Match 0.8%; Score 37; DB 6; Length 937;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 GCTGAACCCGAGAGTGCAGAGAGTCCCGACCTATG 303

DB 198 GCTGAACCCGAGAGTGCAGAGAGTCCCGACCTATG 234

AC ABV58328;
XX
DT ...
XX 13-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 58319.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX PF
XX 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 11198; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 405 BP; 117 A; 98 C; 96 G; 94 T; 0 U; 0 Other;

Query Match 0.6%; Score 30; DB 5; Length 405;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 2378 TTTTTCCTTTTTCTTTCTTTCTTTCTTTCTTTT 2407
Db 93 TTTTTCCTTTTTCTTTCTTTCTTTCTTTCTTTT 64

RESULT 49
ACH20077/c
ID ID ACH20077 standard; cDNA; 408 BP.
XX AC
XX ACH20077;
XX AC
XX DT 13-OCT-2003 (first entry)
XX DE
XX Human adult lung cDNA #1080.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST
KW genome mapping; biodiversity; genetic disorder.

us-10-023-523-17.rng

Wed Jun 9 08:59:02 2004

XX OS Homo sapiens.
XX PN WO200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001339.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-022513P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251983P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX

PA	(HUMA-) HUMAN GENOME SCI INC.
XX	Rosen CA, Barash SC, Ruben SM;
PI	WPI; 2001-465570/50.
XX	Isolated nucleic acid molecule encoding a reproductive system antigen is
DR	used in preventing, treating or ameliorating a medical condition.
XX	Disclosure; SEQ ID NO 9923; 1297pp + Sequence Listing; English.
PT	The present invention provides the protein and coding sequences of a
XX	number of human reproductive system related antigens. These can be used
CC	in the prevention and treatment of reproductive system disorders,
CC	including cancer. The present sequence is a genomic sequence encoding a
CC	protein of the invention
XX	Sequence 2696 BP; 610 A; 697 C; 791 G; 598 T; 0 U; 0 Other;
SQ	Query March 0.6%; Score 30; DB 4; Length 2696;
	Best Local Similarity 100.0%; Pred. No. 0.4;
	Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	2378 TTTTTCCTTTTTCTTTTTCTTTTTTTT 2407
Db	1398 TTTTTCCTTTTTCTTTTTCTTTTTTTT 1427
RESULT 52	
AAL07233	ID AAL07233 standard; DNA; 2696 BP.
XX	AAL07233;
AC	XX
DT	21-NOV-2001 (first entry)
XX	Human reproductive system related antigen DNA SEQ ID NO: 9921.
DE	Human; reproductive system related antigen; reproductive system disorder;
XX	cancer; gene therapy; ds.
KW	Homo sapiens.
KM	WO200155320-A2.
OS	02-AUG-2001.
PN	17-JAN-2001; 2001WO-US001339.
PD	31-JAN-2000; 2000US-0179065P.
PP	04-FEB-2000; 2000US-0180628P.
XX	24-MAR-2000; 2000US-0184664P.
XX	02-MAR-2000; 2000US-0186350P.
XX	16-MAR-2000; 2000US-0189874P.
XX	17-MAR-2000; 2000US-0190076P.
XX	18-APR-2000; 2000US-0198123P.
XX	19-MAY-2000; 2000US-0205515P.
XX	07-JUN-2000; 2000US-0209467P.
XX	28-JUN-2000; 2000US-0214886P.
XX	30-JUN-2000; 2000US-0215135P.
XX	07-JUL-2000; 2000US-0216647P.
XX	07-JUL-2000; 2000US-0216806P.
XX	11-JUL-2000; 2000US-0217487P.
XX	11-JUL-2000; 2000US-0217496P.
XX	14-JUL-2000; 2000US-0218290P.
XX	26-JUL-2000; 2000US-0220963P.
XX	26-JUL-2000; 2000US-0220964P.
XX	14-AUG-2000; 2000US-0224518P.
XX	14-AUG-2000; 2000US-0224519P.
XX	14-AUG-2000; 2000US-0225213P.
XX	14-AUG-2000; 2000US-0225214P.
XX	14-AUG-2000; 2000US-0225266P.
XX	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447F.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0225799P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226868P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	14-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234224P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236367P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT

06-DEC-2000;	2000US-0251479P.
08-DEC-2000;	2000US-0251856P.
08-DEC-2000;	2000US-0251868P.
08-DEC-2000;	2000US-0251869P.
08-DEC-2000;	2000US-0251989P.
08-DEC-2000;	2000US-0251990P.
11-DEC-2000;	2000US-0254097P.
05-JAN-2001;	2001US-02549678P.
(HUMA-) HUMAN GENOME SCI INC.	
Rosen CA, Barash SC, Ruben S	
WPI; 2001-433232/52.	
Nucleic acids encoding 973 hum	
for preventing, diagnosing and	
Disclosure; SEQ ID NO 3433; 75	
The present invention provides	
human testicular antigens, and	
sequences can be used in the t	
reproductive system, immune, i	
gastrointestinal disorders, in	
especially testicular cancers,	
protein fragment of the invent	
Sequence 2696 BP; 610 A; 697 C	
Query Match	0.6%; S
Best Local Similarity	100.0%;
Matches 30; Conservative	0;
2378 TTTTTCCTTTTTCCTTTTTCCT	
1398 TTTTTCCTTTTTCCTTTTTCCT	
RESULT 54	
ABL98793	
ABL98793 standard; DNA; 2696 B	
ABL98783;	
21-JUN-2002 (first entry)	
Human testicular antigen encod	
Human; testicular antigen; tes	
reproductive system disorder;	
cardiovascular disorder; respi	
gastrointestinal disease; infe	
Homo sapiens.	
WO200155317-A2.	
02-AUG-2001.	
17-JAN-2001; 2001WO-US001329.	
31-JAN-2000; 2000US-0179065P.	
04-FEB-2000; 2000US-0180628P.	
24-FEB-2000; 2000US-0184664P.	
02-MAR-2000; 2000US-0186350P.	
16-MAR-2000; 2000US-0189874P.	
17-MAR-2000; 2000US-0190076P.	
18-APR-2000; 2000US-0198123P.	
19-MAY-2000; 2000US-0205151P.	
07-JUN-2000; 2000US-0209467P.	
28-JUN-2000; 2000US-0214886P.	
30-JUN-2000; 2000US-0215135P.	

i testicular antigen polypeptides, useful
 or treating testicular cancer.
 p; English.
 he protein and coding sequences of 973
 fragments of their genomic sequences. The
 treatment of cardiovascular, urinary system,
 spiratory, neurological and
 ceptions, and particularly cancer,
 the present sequence is a DNA encoding a

791 G; 598 T; 0 U; 0 Other;
 re 30; DB 4; Length 2696;
 ed. No. 0.4;
 Mismatches 0; Indels 0; Gaps 0
 TTTT TTTT 2407
 |||||
 TTTT TTTT 1427

G DNA fragment SEQ ID NO: 3435.
 s; cancer; metastasis; immune disorder;
 rinary system disorder; gene therapy;
 tory disorder; neurological disorder;
 ion; cytostatic; gene; ds.

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 11-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.

Disclosure; SEQ ID NO 3435; 766pp; English.

The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system, reproductive system, immune, respiratory, neurological and gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a DNA encoding a protein fragment of the invention

Sequence 2696 BP; 610 A; 697 C; 791 G; 598 T; 0 U; 0 Other;

	Query Match	0.6%; Score 30; DB 4; Length 2696;	
	Best local similarity	100.0%; Pred. No. 0.4;	
	Matches 30; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	2378 TTTTCTTTTTCTTTTCCTTTTTTTT 2407		PR 06-SEP-2000; 2000US-0230437P
			PR 06-SEP-2000; 2000US-0230438P
			PR 08-SEP-2000; 2000US-0231242P
			PR 08-SEP-2000; 2000US-0231243P
			PR 08-SEP-2000; 2000US-0231244P
			PR 08-SEP-2000; 2000US-0231413P
			PR 08-SEP-2000; 2000US-0231414P
			PR 08-SEP-2000; 2000US-0232080P
Db	1398 TTTTCTTTTTCTTTTCCTTTTTTTT 1427		PR 08-SEP-2000; 2000US-0232081P
			PR 12-SEP-2000; 2000US-0231968P
			PR 14-SEP-2000; 2000US-0232397P
			PR 14-SEP-2000; 2000US-0232398P
			PR 14-SEP-2000; 2000US-0232399P
			PR 14-SEP-2000; 2000US-0232400P
			PR 14-SEP-2000; 2000US-0232401P
			PR 14-SEP-2000; 2000US-0233063P
			PR 14-SEP-2000; 2000US-0233064P
			PR 14-SEP-2000; 2000US-0233065P
			PR 21-SEP-2000; 2000US-0234223P
			PR 21-SEP-2000; 2000US-0234274P
			PR 25-SEP-2000; 2000US-0234997P
			PR 25-SEP-2000; 2000US-0234998P
			PR 26-SEP-2000; 2000US-0235484P
			PR 27-SEP-2000; 2000US-0235834P
			PR 27-SEP-2000; 2000US-0235836P
			PR 29-SEP-2000; 2000US-0236327P
			PR 29-SEP-2000; 2000US-0236367P
			PR 29-SEP-2000; 2000US-0236368P
			PR 29-SEP-2000; 2000US-0236369P
			PR 29-SEP-2000; 2000US-0236370P
			PR 02-OCT-2000; 2000US-0236802P
			PR 02-OCT-2000; 2000US-0237037P
			PR 02-OCT-2000; 2000US-0237038P
			PR 02-OCT-2000; 2000US-0237039P
			PR 02-OCT-2000; 2000US-0237043P
			PR 13-OCT-2000; 2000US-0239935P
			PR 13-OCT-2000; 2000US-0239937P
			PR 20-OCT-2000; 2000US-0240960P
			PR 20-OCT-2000; 2000US-0241221P
			PR 20-OCT-2000; 2000US-0241785P
			PR 20-OCT-2000; 2000US-0241786P
			PR 20-OCT-2000; 2000US-0241787P
			PR 20-OCT-2000; 2000US-0241808P
			PR 20-OCT-2000; 2000US-0241809P
			PR 20-OCT-2000; 2000US-0241826P
			PR 01-NOV-2000; 2000US-0244617P
			PR 08-NOV-2000; 2000US-0246474P
			PR 08-NOV-2000; 2000US-0246475P
			PR 08-NOV-2000; 2000US-0246476P
			PR 08-NOV-2000; 2000US-0246477P
			PR 08-NOV-2000; 2000US-0246478P
			PR 08-NOV-2000; 2000US-0246523P
			PR 08-NOV-2000; 2000US-0246524P
			PR 08-NOV-2000; 2000US-0246525P
			PR 08-NOV-2000; 2000US-0246526P
			PR 08-NOV-2000; 2000US-0246527P
			PR 08-NOV-2000; 2000US-0246528P
			PR 08-NOV-2000; 2000US-0246529P
			PR 08-NOV-2000; 2000US-0246610P
			PR 08-NOV-2000; 2000US-0246611P
			PR 08-NOV-2000; 2000US-0246613P
			PR 17-NOV-2000; 2000US-0249207P
			PR 17-NOV-2000; 2000US-0249208P
			PR 17-NOV-2000; 2000US-0249209P
			PR 17-NOV-2000; 2000US-0249210P
			PR 17-NOV-2000; 2000US-0249211P
			PR 17-NOV-2000; 2000US-0249212P
			PR 17-NOV-2000; 2000US-0249213P
			PR 17-NOV-2000; 2000US-0249214P
			PR 17-NOV-2000; 2000US-0249215P
			PR 17-NOV-2000; 2000US-0249216P
			PR 17-NOV-2000; 2000US-0249217P

cystostatic; gene therapy; vaccine; metastasis; ds

Homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US001354.

31-JAN-2000; 2000US-0173065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.
16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198112P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0215647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0228668P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0228287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.

PR	07-JUL-2000;	2000US-0216647P.	PR	20-OCT-2000;	2000US-0241808P.
PR	07-JUL-2000;	2000US-0216880P.	PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000;	2000US-0217487P.	PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225216P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225477P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249285P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249287P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251866P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0

The present sequence is one of 3351 sequences in a library of human polynucleotides. The library is used to detect differentially expressed genes correlated with a cancerous state of a mammalian cell and can detect colon, prostate, breast and lung cancer. The library can be used to produce probes for detection of mRNA and to produce additional copies of the polynucleotides. The probes can be used for chromosome mapping or antisense oligonucleotides for detection of transcription levels. Ribozymes or antisense oligonucleotides can be generated. The polynucleotides and their gene products are used as genetic or biochemical markers (e.g. in blood or tissues) that will detect the earliest changes along the carcinogenesis pathway and/or monitor the efficacy of therapies and preventive interventions. The polynucleotides, polypeptides and antibodies against them can be used in pharmaceutical compositions to

CC genes listed in the specific

XX New cardiovascular system related polynucleotides and polypeptides,
PT useful for preventing, treating, or ameliorating a medical condition,
PT such as cancer of cardiovascular tissues and cancer metastases.
XX Claim 1; SEQ ID NO 2111; 262pp; English.
PS

DT	01-JAN-2004	(first entry)
XX	Human GPCR gene SEQ ID NO:1393.	
DE	ds; gene; human; GPCR;	
KW	guanosine triphosphate-binding protein coupled receptor; gene therapy.	
KX	Homo sapiens.	
OS	EP1270724-A2.	
XX	02-JAN-2003.	
XX	18-JUN-2002; 2002EP-00013517.	
PP	18-JUN-2001; 2001JP-00246789.	
FR	(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.	
PA	(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.	
XX	Suwa M, Asai K, Akiyama Y, Aburatani H;	
PI	WPI; 2003-315783/31..	
XX	P-PSDB; ADC86941.	
XX	New polynucleotide, useful for preparing a composition for treating a	
PT	patient in need of increased or suppressed activity or expression of the	
PT	guanosine triphosphate-binding protein coupled receptor.	
XX	Claim 1; SEQ ID NO 1393; 28pp; English.	
PS	The invention relates to a novel polynucleotide encoding a guanosine	
PS	triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of	
CC	the invention may have a use in gene therapy. The polynucleotide and	
CC	polypeptide are useful for preparing a composition for treating a patient	
CC	in need of increased or suppressed activity or expression of the	
CC	guanosine triphosphate-binding protein coupled receptor. The	
CC	polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the	
CC	invention.	
XX	Sequence 349901 BP; 108973 A; 66040 C; 67419 G; 108968 T; 0 U; 601 Other;	
SQ	Query Match 0.6%; Score 29; DB 9; Length 349901;	
	Best Local Similarity 100.0%; Pred.No. 0.51;	
	Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	2377 CTTTCTTTTTTTCTTTTCCTTTTTT 2405	
DB		
	247310 CTTTTTTTTTTTCTTTTCCTTTTTT 247338	
RESULT 71		
AAL25533		
ID	AAL25533 standard; cDNA; 252 BP.	
XX	AAL25533;	
AC	07-DEC-2001 (first entry)	
XX	Human breast cancer expressed polynucleotide 17990.	
XX	Human; breast cancer; cell marker; cytostatic; ss.	
KW	Homo sapiens.	
OS	WO200151628-A2.	
XX	19-JUL-2001.	
PD	10-JAN-2001; 2001WO-US0000798.	
XX	14-JAN-2000; 2000US-0176077P.	
PR	14-MAR-2000; 2000US-0189167P.	

PR	24-MAR-2000; 2000US-0192099P.	
PR	29-MAR-2000; 2000US-0193480P.	
PR	15-MAY-2000; 2000US-0205230P.	
PR	09-JUN-2000; 2000US-0211315P.	
PR	25-JUL-2000; 2000US-0220534P.	
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
PA	Lillie J, Xu Y, Wang Y, Steinmann K;	
XX	WPI; 2001-451856/48.	
XX	New peptide useful as a marker for the diagnosis of breast cancer.	
PT	Claim 1; Page 3326; 3695pp; English.	
XX	The invention relates to human breast cancer expressed polynucleotides	
CC	(AAL07544-AAL26789) and methods of assessing whether a patient is	
CC	afflicted with breast cancer by examining the correlation between the	
CC	expression of certain markers and the cancerous state of breast cells.	
CC	The polynucleotides and encoded polypeptides are potential markers for	
CC	detecting, diagnosing, monitoring, characterizing treating and	
CC	potentially preventing breast cancer. The polynucleotides and encoded	
CC	polypeptides are also useful for isolating compounds with cytostatic	
CC	activity	
XX	Sequence 252 BP; 75 A; 31 C; 40 G; 106 T; 0 U; 0 Other;	
SQ	Query Match 0.6%; Score 28; DB 4; Length 252;	
	Best Local Similarity 100.0%; Pred.No.2.9;	
	Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	2369 TCTCTTTCTCTTTCTTTCTTTCTTTCTTTT 2396	
Db	64 TCTCTTTCTCTTTCTTTCTTTCTTTCTTTT 91	
RESULT 72		
AAL16690		
ID	AAL16690 standard; cDNA; 265 BP.	
XX	AC	
XX	AAL16690;	
XX	07-DEC-2001 (first entry)	
DT		
XX	Human breast cancer expressed polynucleotide 9147.	
DE		
XX	Human; breast cancer; cell marker; cytostatic; ss.	
XX	Homo sapiens.	
OS		
XX	WO200151628-A2.	
PN		
XX	19-JUL-2001.	
PD		
XX	10-JAN-2001; 2001WO-US000798.	
PF		
XX	14-JAN-2000; 2000US-0176077P.	
PR	14-MAR-2000; 2000US-0189167P.	
PR	24-MAR-2000; 2000US-0192099P.	
PR	29-MAR-2000; 2000US-0193480P.	
PR	15-MAY-2000; 2000US-0205230P.	
PR	09-JUN-2000; 2000US-0211315P.	
PR	25-JUL-2000; 2000US-0220534P.	
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
PA	Lillie J, Xu Y, Wang Y, Steinmann K;	
XX	WPI; 2001-451856/48.	
XX	New peptide useful as a marker for the diagnosis of breast cancer.	
PT		
XX		

1

The invention relates to human breast cancer expressed polynucleotides (AAL075x4-AAL26789) and methods of assessing whether a patient is

XX	WPI; 2001-562795/76.
XX	
XX	Novel isolated nucleic acid molecule associated with cancerous state of
PT	prostate cells and correlating with presence of prostate cancer, useful
PT	for detecting presence of prostate cancer, stage of prostate cancer.
XX	
XX	Claim 1; Page 4141; 11750pp; English.
PS	
PS	
CC	The invention relates to an isolated nucleic acid molecule (I) comprising
CC	a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC	specification or its complement. (I) is useful for: (a) assessing whether
CC	a patient is afflicted with prostate cancer; (b) monitoring the
CC	progression of prostate cancer in a patient; (c) assessing the efficacy
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC	determining whether prostate cancer has metastasized in a patient; (h)
CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX	
XX	Sequence 2178 BP; 502 A; 544 C; 457 G; 671 T; 0 U; 4 Other
SQ	

Query Match	0.6%;	Score 28;	DB 5;	Length 2178;
Best Local Similarity	100.0%;	Pred. No. 2.2;		
Matches 28;	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy 2379 TTTTTTTTTTCTTTTCTTTTTTTT 2406
 |||||
 Db 660 TTTTTTTTTTCTTTTCTTTTTTTT 633

RESULT 86

ABV28966/c
ID ABV28966 standard: cDNA: 2178 BP.

XX ABV28966:

XX
DT 16-SEP-2002 (first entry)XX
DE Human prostate expression marker cDNA 28957.

xx Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
kw pharmacogenomic marker; gene: ss.
kw

XX
05 Homo sapiens

XX PN WO200160860-A2

XX
PD
23-AUG-2001XX
PF 20-FEB-2001: 2001WO-US005171.XX
PR 17-FEB-2000: 2000US-0183319P.

PR 16-MAR-2000; 2000US-01899862P.
PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P

PR 13-DEC-2000; 2000US-0255281P.
YY

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

XX
DR WPI: 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer. stage of prostate cancer.

xx
ps
Claim 1: Page 6121-6122: 11750pp: English.

XX 25

[illegible]

RESULT 87
ABV28140/C

ID ABV28140 standard; cDNA: 2178 BP.

XX
AC : ABV28140:

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 28131.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene: ss.
KW

OS Homo sapiens.

XX
PN
WO200160860-A2.

23-AUG-2001.

AA
PF 20-FEB-2001; 2001WO-US005171

XX
PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-018986ZF.
PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

WPI: 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS
Claim 1: page 5830-5831: 11750pp; English:

xx The invention relates to an isolated nucleic acid molecule (I) comprising
cc a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
cc specification or its complement. (I) is useful for: (a) assessing whether
cc a patient is afflicted with prostate cancer; (b) monitoring the
cc progression of prostate cancer in a patient; (c) assessing the efficacy
cc of a test compound to inhibit prostate cancer in a patient; (d) assessing
cc the efficacy of a therapy for inhibiting prostate cancer in a patient;
cc (e) selecting a composition for inhibiting prostate cancer in a patient;
cc (f) assessing the prostate cell carcinogenic potential of a compound; (g)

AAAC74446 to AAC77606 encode the proteins given in ABA540237 to ABA543397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticongestant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; CC

products. AAH51627 - AAH51631 and AAB62916 - AAB62918 represent g35018 cDNA sequences and protein products. Primers AAH51632 - AAH51699 are used to isolate sbg1 cDNAs, while sbg1 exons from different primates are represented by sequences AAH51642 - AAH51699. Nucleotide sequences of amplicons which comprise biallelic markers located on the chromosome 13q31-q33 locus are represented in AAH51700 - AAH51817. Biallelic markers are represented in the sequences by degenerate/undefined base codes. PCR primers AAH51818 and AAH51819 are used in the isolation of sequences of the invention. The biallelic marker containing nucleotide sequences are used to determine the identity of the nucleotide at a biallelic marker in a sample DNA sequence. The nucleotide sequences may be labelled and used for genotyping by determining the identity of a nucleotide at a Region D-related biallelic marker in a biological sample from single or multiple subjects. By determining the frequency of a biallelic marker in a population an association between a genotype and a trait, a haplotype and a trait and a phenotype and a trait can be detected. The sequences can be used to determine a predisposition to or early onset of schizophrenia or bipolar disorder or a beneficial response to or side effects related to treatment against schizophrenia or bipolar disorder.

XX	Sequence	3001	BP;	824	A;	674	C;	534	G;	965	T;	0	U;	4	Other;		
SQ	Query Match									0.6%;	Score	28;	DB	3;	Length	3001;	
	Best Local Similarity									100.0%;	Pred.	No.	2.1;				
	Matches	28;	Conservative							0;	Mismatches				0;	Gaps	0;

[illegible]

RESULT 90
ACF05254
ID: ACF05254 standard: cdna: 3197 BP.

XX	ACF05254;
XX	
AC	
XX	.06-NOV-2003 (first entry)
DT	
XX	
XX	Plasmodium falciparum cGMP dependent protein kinase cDNA.
DE	
XX	
XX	Protein kinase; enzyme; malaria; vaccine; diagnosis; therapy;
KW	protozoacide; gene; ss.
KW	
XX	
OS	Plasmodium falciparum.

xx	Key	Location/Qualifiers
xx	FH	101..2662
xx	FT	/*tag= a
xx	FT	/product= "cGMP-dependent protein kinase"
xx	FT	

AA
PN
WO2003054157-A2XX
PD 03-JUL-2003.

18-DEC-2002: 2002WO-US040527.

20-DEC-2001: 2001US-0342816P.

PA (MERI) MERCK & CO INC.

PI Liberator P, Schmatz D, Gurnett A, Diaz C;

XX
DR WPI: 2003-559137/52.

DK F-FSDB; ABR02490.
XX

AA New isolated or recombinant nucleic acid molecule encoding Eimeria maxima
PT or Plasmodium falciparum cyclic GMP-dependent protein kinase, useful for
PT diagnosing, preventing or treating protozoal infections.

PS · Claim 26; Page 54-57; 29pp; English.

XX CC The present sequence is that of a cDNA clone encoding a novel cGMP-

CC dependent protein kinase (pK) of plasmodium falciparum. The clone was
CC obtained from P. falciparum trophozoite cDNA by PCR using primers based
CC on cAMP-binding and catalytic domain peptides (see ABR62497-501). PKG
CC nucleic acids, polypeptides, antibodies and inhibitors can be used in the
CC diagnosis of infection, as vaccines, as chemotherapeutic targets, and in
CC screening assays to identify potential antiprotozoal agents
XX
XX Sequence 3197 BP: 1312 A; 301 C; 496 G; 1088 T; 0 U; 0 Other;

Query Match	0.63;	Score 28;	DB 8;	Length 3197;
Best Local Similarity	100.0%;	Pred. No. 2.1;		
Matches 28;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	

[illegible]

RESULT 91
ABN83030/c
ID ABN83030 standard: CDNA: 3905 BP.

AA
AC ABN83030:

14-AUG-2002 (first entry)

Human X2 theca repressor 10.34 cdna.

Human; X2 theca repressor 10.34; repressor; immunopathy; gene; ss.

OS Homo sapiens.

Key	Location/Qualifiers
FH	

```
FT
/*tag= a
```

XXIX

XX
X

XX

XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

X

DR P-PSDB; ABB81691.

PT Polypeptide-human X2 theca repressor 10.34 and polynucleotide for coding
PT it.

PS Claim 6; Page 25-27; 34pp; Chinese.

CC The sequence encodes the novel human X2 theca repressor 10.34 polypeptide
CC of the invention. The polypeptide is useful for treating diseases such as
CC immunopathy. The antagonist of the polypeptide and its medical action,
CC and the application of the polynucleotide are also disclosed

Sequence 3905. BP: 1245 A; 731 C; 701 G; 1228 T; 0 U; 0 Other; XX SO

. Query Match 0.6%; Score 28; DB 6; Length 3905;

```
..Base local similarity score: 0.97; local alignment: 100%
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

2369 TTCCTTTTCTTTT 2396

D_B

3466 TTCTTTTCTTTTTTTTTTCCTTTT 3439

[illegible]

